IMPLEMENTATION OF PERFORMANCE MANAGEMENT IN REGIONAL GOVERNMENT IN RUSSIA

by

Alexander Kalgin

A thesis submitted to the University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

University of Birmingham

International Development Department

December 2014

ABSTRACT

The thesis explores the experience of implementing a system of performance management in regional governments in Russia. The findings suggest that the system was affected by deliberate data manipulation. Using mixed methods I demonstrate that locally generated data were more likely to be manipulated than the data reported by external agencies. Instead of improving managerial decisions, performance indicators have become a tool of symbolic bureaucratic accountability weakly linked to real managerial activities. 25 current and former civil servants from three regional governments in Russia were interviewed (including three ministers of economic development); quantitative data were obtained from a publicly available performance dataset covering the period of 2007-2011 (with data for a unified list of over 300 indicators from 83 regional governments). Two strategies of data manipulation were identified: a "prudent bureaucrat" strategy consisted in minimizing long-term risks by reporting "more-normal-than-real" figures; a more ambitions "reckless bureaucrat" strategy aimed at inflating figures to maximise credit. Systematic application of these two strategies has produced a detectable bias in the overall performance data with "prudent bureaucrat" strategy dominating. A survey of 170 municipalities was carried out to demonstrate that data for indicators most affected by "prudent" manipulation are perceived as less trustworthy by civil servants. I then theorize that performance reporting creates a "bureaucratic panopticon" and resulting behaviour may be interpreted using Michel Foucault's notion of normalisation and Herbert Simon's model of "administrative man".

ACKNOWLEDGEMENTS

I would like to thank the Higher School of Economics for the scholarship that made this thesis possible. I would particularly like to express my gratitude to my co-supervisor Dr. Andrey Klimenko for his support and initiative on early stages of preparation for this research project.

I would like to thank the staff of the International Development Department for their flexibility and patience in administering all the necessary documents that made this research possible.

I would like to thank Dr. Michael Hubbard for his encouragement of this work.

I would like to thank my supervisors – Dr. Adrian Campbell and Dr. Andrey Klimenko for their help and support during the process of preparing this thesis.

I would also like to thank my parents for their understanding and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

DROND – *Doklad o Rezultatakh i Osnovnykh Napravleniakh Deyatelnosti* (rus. Report on Performance and Future Actions)

FSO – *Federalnaya Sluzhba Okhrany* (rus. Federal Security Service)

NPM – New Public Management

PM – Performance Management/Measurement

Rosstat – *Federalnaya Sluzhba Statistiki* (rus. Federal Bureau of Statistics)

CHAPTER 1. INTRODUCTION

Performance management remains a popular instrument of public sector reform. Russian government has introduced a system of performance management at the regional level. This system utilized the inventory of over 300 indicators to monitor the performance of regional administrations in all 83 Russian regions. The system has been widely criticized by both civil servants and academics. It has been described as overly cumbersome and having little managerial value. This work sets out to explore the experience of civil servants in dealing with the system, to describe problems associated with it and to identify managerial gains produced by the system. Particular attention is drawn to the problem of data manipulation. It is shown that traces of manipulative behaviour are discernible in the data generated by the system. Qualitative and quantitative data are triangulated to provide a more vivid picture of the system.

The practice of performance management has been under strong criticism from some scholars who pointed at multiple unintended consequences arising from the practice of formal performance measurement (Bevan and Hood 2006; Bohte and Meier 2000; Hood 1991, 2006; Patrick 2009; Smith 1995). Russian commentators have particularly stressed the administrative burden of operating a system with an excessive number of indicators. Other criticisms included lack of feedback, insufficient transparency in setting performance indicators (Malkov 2010) and inadequate coordination with other reform efforts in the public sector (Verheijen and Dobrolyubova 2007).

On the positive side, respondents during interviews indicated some managerial gains derived from the system. It has been noted that the system has become a tool facilitating inter-regional competition through benchmarking and was instrumental in identifying areas of governmental failure in providing services.

Data manipulation was selected as an area for further exploration. Both qualitative and quantitative evidence of deliberate distortion of performance figures were obtained. There is a

large body of literature on detecting deliberate data manipulation in the private sector (Beatty et al. 2002; Burgstahler and Dichev 1997; Dechow et al. 1995; Degeorge et al. 1999; J. J. Jones 1991), in contrast, empirical literature on public sector performance measurement is relatively scarce (Charbonneau and Bellavance 2012; Gueorguieva et al. 2009). There have been a number of literature reviews and conceptual papers (Hood 1991, 2006), but not many empirical studies.

In the Russian context the system of nation-wide performance measurement was introduced together with other significant public sector innovations as part of a comprehensive reform package known as "the Administrative reform" in 2007. Since then it has seen a number of revisions and is currently still in operation. The implementation process revealed significant obstacles to effective functioning of performance measurement. Issues of interagency cooperation and public sector integrity became troublesome. Performance measures were introduced in the context of a highly centralized public sector hierarchy and became integrated in the informal system of incentives of the public sector. The functioning of performance management was thwarted by lack of managerial autonomy. Observing the results of the process of implementation reveals inherent problems of the system of performance management in Russian regional governments.

Research objective and questions

Although the implementation of performance management in western contexts is well researched, its implementation in developing countries is relatively scarcely covered by the literature. The objective of this thesis is to demonstrate how performance indicators function in a specific institutional environment of Russian regional governments. To achieve this objective one particular aspect of the working of the performance management system was selected. Data manipulation practices were chosen for deeper investigation.

The following research questions were formulated:

1. How do formal performance indicators function in the context of the Russian regional government?

2. What managerial benefits do civil servants see in the system?
3. Is the system of performance management susceptible to deliberate data manipulation?
4. If yes, then what type of data manipulation is most common?

Research methods

The research is conducted using mixed methods approach. A combination of quantitative and qualitative methods is used. Three regional governments were visited to conduct interviews and observe the practice of regional civil servants. A variety of methods of data collection was used: literature reviews, semi-structured interviews, questionnaire surveys and participant observations.

The research was conducted in three stages: first, 25 semi-structured interviews were carried out with officials from regional governments, consultants and academics[1]. Motives and opportunities for deliberate data manipulation were identified and hypotheses regarding their manifestation in resulting performance data were formulated. Second, quantitative evidence was obtained by analysing the nation-wide performance dataset[2] that provides figures for over 300 performance indicators covering 83 regional governments for the period of 5 years between 2007 and 2011 (this gave over 90 000 observations). Third, a survey of municipal civil servants was conducted to test the predictions of the quantitative analysis.

[1] The sample included: 16 current regional civil servants, including three ministers of economic development, three deputy ministers from departments of health and economic development, 6 heads of departments and 2 specialists (lower level civil servants); 9 former civil servants, including: 3 consultants, 3 academics. Other respondents included: a top central government official responsible for the design of the nation-wide system of performance management and two civic activists involved in public scrutiny of government performance reports. The interview guide from Christopher Pollitt, 'Unleashing Change: A Study of Organizational Renewal in Government', *Administrative Science Quarterly*, 51/2 (2006b), 302-04. was used with minor alterations. Purpose sampling and snowballing were used to generate the list of respondents. People responsible for implementing and operating performance measurement systems were identified in regional governments and asked to provide further contacts for interviews.

[2] The data were collected in accordance with the Presidential Decree № 825 of 28.06.2007 "On assessing the effectiveness of executive authorities of the subjects of the Russian Federation.", The data were published by the Ministry of regional development in 2012.
URL: http://www.minregion.ru/upload/documents/2012/10/101012/101012_itogi_2011.xls [accessed 08.03.2013]

Structure of the thesis

The thesis is divided into nine chapters including the introduction (Chapter 1).

Chapter 2 outlines the background of performance management reform in the Russian public sector. The chapter provides background information on the history of public sector reform in Russia, gives an overview of the structure of Russian government and discusses the specifics of the Russian usage of performance management terminology.

Chapter 3 reviews the literature on 1) theories of bureaucracy, 2) performance management and New Public Management (NPM), 3) unintended consequences of implementing performance management regimes; 4) data manipulation in private and public sector contexts, and 5) underlying administrative values of public sector reforms.

Chapter 4 describes methods of the research and outlines epistemological and ontological assumptions.

Chapter 5 provides qualitative findings. This chapter outlines the problems in implementing performance measures as perceived by civil servants. The following issues are identified: the inability of agencies to influence performance indicators, the importance of external factors, lack of interagency cooperation, the importance of informal functions of performance reporting, low trust in performance data.

Chapter 6 presents quantitative findings. The problem of data manipulation is investigated using statistical methods. Two strategies of data manipulation are identified, hypotheses regarding expected patterns in performance data are formulated and tested. The results of the survey of municipal civil servants are also presented in this chapter and triangulated with the rest of the findings.

Chapter 7 links the results of the research to wider theoretical discussions. Theoretical generalizations are made using theories of bureaucracy of Gordon Tullock and Antony Downs. The results are interpreted using theories of bounded rationality (Herbert Simon), normalization (Michel Foucault).

Chapter 8 discusses alternative interpretations of the role of performance measurement and uses the theory of ideology (Slavoj Zizek) to give an alternative interpretation to the findings. Limitations of the research are discussed.

Chapter 9 concludes by summarizing the answers to research questions.

Two narratives

The thesis is built around two narratives that capture significant aspects of practices related to performance reporting.

The first narrative is structured around the process of implementation and operation of several waves of performance measurement/management reform initiatives. This narrative focuses on day-to-day experience of civil servants and captures different views on the system held by civil servants at different levels of bureaucratic hierarchy: views of regional civil servants, central government officials, middle management and senior civil servants. This narrative captures the perception of performance measurement from within the bureaucratic hierarchy. Using Hegelian terms it may be said that it attempts to demonstrate what performance measurement practices are *for-itself*: how they are perceived by those who take part in them and practice them. This narrative is written in the "understanding" mode. Such issues as effectiveness, cooperation, autonomy, managerial discretion are explored in the context of regional governments.

The second narrative is structured around the problem of data manipulation and demonstrates that reported performance figures may suffer from deliberate distortions. Theoretical explanation is constructed using theories of bureaucracy to show that such distortions may result from the desire of civil servants to avoid attention of their superiors. Explanation proceeds using agency and public choice theories and such theoretical conjunctures as "bureaucratic personality types", "self-interested behaviour" etc. In these models the ultimate given is a model of man and this man's motives and drives. The situation is analyzed to the point where an explanation may be made using basic ultimate givens such as self-interested motives of

bureaucrats. These are, of course, conjunctures of the researcher – abstract constructions whose aim is to simplify the complexity of the real world and render it intelligible and predictable. By using them it is not implied that civil servants think of themselves as self-interested rent-optimizers, but it is merely asserted that such simplified models are useful in deriving an explanation of the situation that allows for predictions that approximate observed data.

What performance reporting practices are *for-themselves* depends on what they are *in-themselves* (and vice-versa). This may be illustrated using the following hypothetic construction: if the civil servant's main purpose is to diligently serve the public good then the system of performance reporting would be used and perceived as a tool making this purpose more or less easily fulfilled. If, on the other hand, their main aim is to extract rent and seek bribes, then the same system would be perceived and (ab)used accordingly.

The story of performance management in the Russian public sector presented in this work is at times critical. The Russian practice demonstrated that performance management could be either useful or wasteful depending on a particular case. Russia here is not exceptional. Researchers of the British experience have also found failures and inefficiencies in performance management regimes in the UK (Pidd 2005; Public Administration Select Committee 2003; Radnor 2008). Russia is not exceptionally bad, neither is it exceptionally good in terms of implementation of PM. The account presented below should be read with this in mind. The Russian experience is not measured against a practical international standard achieved elsewhere. It is rather measured against theoretically achievable ideals or commonly shared expectations and, for this reason, may sometimes be too skeptical. Practice inevitably falls short of theoretical ideal conjunctures.

CHAPTER 2. PUBLIC SECTOR REFORMS IN RUSSIA

Historical background

> *Soviet legacy accounts for much of the current character of the public service. This legacy also explains why Western experiences and models are unlikely to be grafted wholesale onto the Russian administrative landscape. (Barabashev and Straussman 2007: 373)*

In analyzing public sector reforms in modern Russia it is important to consider the historical background that shaped modern Russian society and state. It is impossible to draw a line in the past that would set the boundary for the analysis of the relevant historical context. Events of the ancient history influence our life up to the present day. One can hardly mark a point in the past where relevant trends begin. The traditions of ancient Kiev Russia affected the outcome of the reforms of Peter the Great; the Czarist Russian Empire determined the path of the development of the Soviet Union; the Soviet legacy continues to live in the modern institutions of post-Soviet Russia. Institutes are slow to change and hard to remodel. It is, therefore, a matter of practicality and limitations of the present analysis to choose the starting point for the discussion.

It would be reasonable to argue that the most dramatic changes in the structure of the Russian society happened during the Soviet era. While preceding history of the Russian Empire inevitably determined some of the characteristics of the modern Russian state (such as its geographical characteristics, for example), its influence on the modern society can only be viewed through the prism of the Soviet era. It is beyond the scope of this study to cover the history of reforms of bureaucracy in Russia and the Soviet Union. There is a whole body of "sovietological" literature (see, for example (Chekharin 1977; Nove 1965; Piekalkiewicz and Hamilton 1991; Piskotin and Sayer 1989). This section is limited to post-1990 trends in the history of the Russian public sector.

This history may be crudely broken down into two stages: early Russian bureaucracy (1990-2000), and modern Russian bureaucracy (2000-present). For the needs of the present

analysis such crude categorization will suffice. Only a brief outline of modern history is given below for the sake of introducing the reader to the Russian present day situation.

Bureaucracy in the Russian federation

> *The 1990s were not kind to public service reform in Russia. (Barabashev and Straussman 2007: 379)*

The history of public sector reforms in modern Russia may be subdivided into 5 distinct stages. Barabashev and Straussman (2007) identify 4 stages in the period between 1992 and 2006 and Jakobson (2001) adds another transition stage in 1990-1992.

These stages include: the first stage of turbulence immediately after the collapse of the Soviet Union, when the raging political struggle meant that the executive functioned without well defined and certain leadership (1990-1991) (Jakobson 2001); the three following preliminary stages that set the scene for a systematic reform effort (1992-2003); the contemporary stage when a relatively consistent reform process began (2003-present) (Barabashev and Straussman 2007).

The zero stage: the big bang

> *Although some administrative reforms were tried repeatedly in 1990–1993, they gave way to an open struggle between irreconcilable political forces, each of which claimed to be legitimate. The question was no longer how to improve the operational performance of the government, but what should be the structure and the principal goals of the state. (Jakobson 2001: 31)*

The political turbulence that followed the collapse of the Soviet state made no systematic reform of the public sector possible during the first two-three years after the event. In fact, the bureaucracy during this period lacked well-defined consistent leadership and had to adapt to rapid changes. This was the time of unprecedented uncertainty, dramatic changes and instability:

> *During that period, Russia twice survived situations in which a universally-acknowledged power system was practically absent: In 1991, when Gorbachev's administration of the Soviet Union and Yeltsin's administration of the Russian Federation began to compete; and in 1993, when President Yeltsin openly fought*

> *against the Supreme Soviet (the parliament) and forbade the officials to execute its laws. (Jakobson 2001: 31)*

Jakobson (2001: 31) remarks that adaptation mechamisms of the bureaucracy during the period of constitutional crisis are in their own right of high interest for scholars of public administration.

The end of this stage is marked by the adoption of the new Constitution of the Russian Federation that established the division of powers; defined the scope of the public sector and distinguished between levels of public administration.

The new constitution established the country as a federation and defined the division of powers. This set the foundation for further refinements in the definition of what constitutes the public sector in Russia. The modern definition of the public sector includes later modifications by the legislature, but at this stage the general design of the Russian public sector was outlined (Barabashev and Straussman 2007: 374). Executive, legislative and judicial powers were assigned their responsibilities; federal, regional and municipal levels of government received their definitions and responsibilities. The municipal level was excluded from the definition of the public sector. This provided the starting point for systematic reforms of the public sector (Barabashev and Klimenko 2010).

The first state: early attempts of systematic reform

The first systematic attempt to reform the public sector for the needs of the new post-communist state began in 1992-1993 (Barabashev and Straussman 2007). The most urgent necessity was to create a system of recruitment, training and promotion of civil servants that did not hinge on the Communist Party apparatus. It was necessary to fill the place of the communist party with something ideologically neutral and technocratic:

> *The first step was the establishment of a specialized public institution responsible for all administrative reform service issues — the Roskadri of Russia. The Roskadri had three goals: (1) to increase the level of education of government employees, (2) to create a system of public service administration modelled partly on Western standards, and (3)*

> *to evaluate the professional preparation of public employees. However, only the first goal was accomplished.*

In the soviet period the communist party was responsible for supervising the public sector. Recruitment, promotion and evaluation of public servants were performed by the party apparatus. Members of this apparatus were bound together by ideology.

The party performed the function of the "second-order cybernetic loop" (Smith 1995: 300) in the system and exercised the control over the "first-order cybernetic loop", that is, the bureaucracy (Institute of the national project «Obshestvenniy dogovor» 2011). In a democratic state the second-order cybernetic loop functions through the institutes of representative democracy and is (in theory, at least) ultimately powered and directed by the will of the people. The history of the reforms of the public sector in Russia up to the present day has been the history of attempts to bring the bureaucracy under the control of the people:

> *… it is well known that the soviet system of government had two managerial circuits. The bureaucracy per se was under the control of the party apparatus: one bureaucratic system was controlled by another. The major change that occurred in 1980s-90s was the removal of this second circuit of control. The Yeltsin era was the time when the society was adapting to the new system. It was attempted to replace the control of the party by democratic institutions of citizen control. (Institute of the national project «Obshestvenniy dogovor» 2011: 20)*

The peculiar characteristic of a soviet bureaucrat is that he is not politically neutral. Unlike in the Western tradition, where political appointments are only made at the very top of the hierarchy, in the Soviet union all managerial positions were held by members of the communist party. This meant that they had to abide not only by the formal rules and regulations (which are easy to change), but also by informal rules and norms, associated with the communist doctrine. Thus, changing the written rules would not make much difference in such a system, if old personnel remained in key positions:

> *…the basic rules of the Soviet system presupposed that officials would have a personal devotion to certain ideas. Practically all the positions responsible for making significant decisions could be occupied only by Communist Party members. According to the Rules of CPSU (Communist Party of the Soviet Union), its members had "to master Marxist-*

Leninist theory, to lead a resolute struggle with the manifestations of bourgeois ideology, of private property psychology, of religion, etc." (Ustav Kommunisticheskoi partii Sovetskogo Souza, 1976), (Jakobson 2001: 32).

In the ideal Weberian-style bureaucracy it would not matter much if old people were in charge or new ones. If the rules were changed such ideal bureaucrats would start working by the new rules:

The prevalence of the old staff would not make so much difference if this type of Soviet official resembled the Weberian bureaucrat. If the officials drew a clear distinction between government service and private life and were guided by comprehensive formal rules, changing these rules could have been compared to setting up a new programme on a computer (Jakobson 2001: 32).

But, since there were no such Weberian bureaucrats, it was necessary to kick-start a process of replacement of the old staff by new employees. This was attempted during this early stage, but with humble success:

…the efforts of the Roskadri were resisted by government employees who did not want to be evaluated, and the agency failed to achieve the political traction in the governmental hierarchy necessary to have any major impact. As a result, the second and third goals remained unfulfilled (Barabashev and Straussman 2007: 378).

In 1995 a federal law[3] was passed that introduced "the concept of public service and defined how it was supposed to function" (Barabashev and Straussman 2007: 378). However, "no clear goals of administrative reform were articulated, thus producing inconsistent interpretations of the law and several incompatible directions for development of a modern public service" (Barabashev and Straussman 2007: 378).

The second stage: adopting new principles

Attempts to make the bureaucracy serve the public interest continued in 1997-98:

The second reform effort was undertaken in 1997 – 98 by a group of advisors to President Boris Yeltsin within the framework of the Conception of Public Administration Reform. Borrowing from international practice, this eff ort tried to identify key elements of a modern public service, such as the principles embodied in the merit system and the

[3] No. 119-FZ, July 31, 1995.

> *notion that public servants serve the public interest. (Barabashev and Straussman 2007: 378)*

These principles were in sharp contrast with the Soviet tradition. The interests of mere citizens were never considered a worthy goal in the past:

> *Soviet bureaucrat's job was to implement and even predict the leadership's demands and to follow the party guidelines, but never to work directly in the interest of citizens — a bourgeois concept, at best (Barabashev and Straussman 2007: 376).*

Thus, an attempt was made to "install new software" into the bureaucratic computer. It was achieved without significant resistance, but did not lead to a real change in the working of the bureaucracy:

> *The most one could say about [these] efforts …is that they led to the adoption of selected principles by both professionals and political officials without significant resistance; however, no mechanisms for the implementation of these principles were established. The reform gradually terminated without ever really getting off the ground. (Barabashev and Straussman 2007: 376).*

Such ideological flexibility is characteristic of the particular kind of bureaucratic mentality that evolved during the soviet era. The core feature of this mentality is the acceptance of the fact that political declarations and formal written rules have nothing to do with the real life. Thus, fashionable buzzwords may come and go with no impact on actual behaviour. This became obvious during the process of disintegration of the Soviet Union, when communist party members demonstrated only slight resistance to the change despite being "true communist" on paper and in public speeches:

> *Although the public administration was dominated by members of CPSU, a shift away from communism was accomplished mainly by the most active part of the staff, the majority being neutral and a minority being slightly opposed (Jakobson 2001: 33).*

Jakobson (2001: 33) asserts that this may be explained by distinguishing between formal and informal rules adhered by the bureaucrats. Three characteristics of soviet bureaucratic mentality may be identified:

(1) a profound difference between formal requirements (fixed in written form) and real life;

(2) the ideologization of formal rules, which means that terms like 'communism' and 'plan' could easily be replaced by the words 'democracy, 'market,' etc., since they are only a kind of formula;

(3) the strong influence of certain informal rules that do not presuppose the devotion to a set of abstract ideas, including the idea of honest performance of official duties (Jakobson 2001: 33).

This mentality was preserved by the staff of the new Russian state as the turnover of civil servants (especially at the middle level and in the regions) was slow:

[By 2000] over 40% of Russian officials used to work in the bureaucracy of the Soviet Union. New staff members were constantly hired in the Soviet era, and very few employees got fired. In three senior groups, officials with over 15 years' experience in civil service constitute approximately half of all current employees (54%, 50%, and 38%, respectively). These employees were promoted during pre-Gorbachev times. There are many high-ranking former Soviet officials among the ministers and their deputies. New civil servants more often hold lower positions in the hierarchy. In Russia's regional governments, the climate is more resistant to change since in most cases the authorities hold less reform-oriented views than federal administrators (Jakobson 2001: 33).

In sum, new principles were declared and accepted but the process of actual implementation did not yet start during this stage.

The third stage: laying the legal foundations

During this stage the legal foundations for implementing public sector reforms were laid. Four pieces of legislation were adopted:

· The Conception of the State Service System Reform of the Russian Federation (Conception);
· About the System of State Service of the Russian Federation (System);
· About the Civil Service of the Russian Federation (Civil Service);
· State Service Reform in the Russian Federation: 2003 – 2005 (Federal Program) (Barabashev and Straussman 2007: 378).

The chief achievement during this stage was that the structure of the public sector was clearly defined:

> *[The legislation] established three main categories of public servants (federal public service, military service, and law enforcement service) and two levels of employees (federal public service and regional public service). (Barabashev and Straussman 2007: 378).*

At this stage the concepts that were outlined in the constitution were worked into detailed administrative regulations and provided the starting point for rational systematic further improvement:

> *…the legislation established the administrative preconditions to implement all of these reforms at both the federal and regional levels. (Barabashev and Straussman 2007: 378).*

The fourth stage - contemporary period: process-control vs managerial autonomy

It is useful here to consider a wider context of public sector reform in Russia to be able to assess the difficulties of implementing performance management/measurement.

In 2003 all initiatives in reforming the public sector were grouped into two large programmes: the *Administrative reform* and the *Civil service reform*. The first group included all initiatives aimed at changing methods of management: budgeting, planning, performance management, etc. The second group included initiatives concerned with human resource management: recruitment, salaries, evaluation mechanisms, etc. The two reforms are currently being implemented as distinguishable programmes.

The design and implementation of public sector reforms posed a serious challenge. Parts of the reform agenda were allocated to different central government entities. As a result, **three distinct reform strands** developed*: the civil service reform*, *the budget reform* and *the administrative reform* (Verheijen and Dobrolyubova 2007: 210). The civil service reform aims at developing a professional civil service by creating incentives for performance at the individual level. The budget reform aimed at improving the budget process and inter-budgetary relations (starting as a separate effort in 2004). The administrative reform aimed at optimising the structure of government and streamlining its operations

Performance management initiatives have been mainly implemented within the framework of the administrative reform. Commentators observe that the principles of performance measurement were often not adequately reflected in the two remaining strands of reforms (Verheijen and Dobrolyubova 2007). Performance measurement mechanisms in inter-governmental relations were not synchronized with corresponding initiatives in performance budgeting or individual performance-related pay.

The major achievement of the administrative reform has been the introduction of a new three-tier government structure at the federal level (Ministry-Service-Agency). This setup aimed at separating the function of policy development from functions of policy implementation, oversight, control and property management. At the federal level this reform became institutionalised through the adoption of the Administrative Reform Concept in 2005, which declared performance management a focal point for improving service delivery and decreasing administrative barriers in the country (Verheijen and Dobrolyubova 2007: 210). Another major direction of the administrative reform was related to streamlining government operations by introducing "administrative protocols" (*административный регламент*): formalized process descriptions of every function and service delivered by the government. It is useful to consider other elements of the administrative reform that are being implemented in parallel with performance management/measurement.

Administrative reform package: contradictory intentions

Since its inception in 2003 the administrative reform has by now gone through three cycles[4]. **The first cycle** began in 2003 and included two main activities:

1. Inventory of all functions of federal and regional government bodies and elimination of duplicating functions;
2. Creation of the threefold system of ministry, service and agency at the federal level and delineation of responsibilities between them;

[4] http://ar.gov.ru/about/history/

3. Clearer delineation of responsibilities between federal and regional authorities.

The general technocratic aim behind these efforts was to formalize the list of public sector functions (services) and assign responsibilities to different types of public bodies. It was believed that one of the key causes of ineffectiveness was high discretion given to bureaucrats at low-to-middle level. This discretion, presumably, allowed for rent-seeking and arbitrary decision making. Thus, once the list of functions was formalized it was possible to begin the process of assigning functions to levels of government, formalizing the process of performing a function or providing a service. This process of *administrative regulation* (formalization) has been the leitmotiv of the administrative reform up to the present day (Klimenko, 2012)

The second cycle began in 2005, when the Conception of the Administrative Reform was passed. The conception was subdivided into 7 reform initiatives:

1. Performance management and project management;
2. Administrative regulation and standardization of public (municipal) functions and services;
3. Providing electronic access to public services;
4. Streamlining public services through “one-window” schemes and “multi-functional centres”;
5. Optimization of government functions and services;
6. Formalizing public sector procurement;
7. Reducing corruption (Concept of Administrative Reform 2005-2010).

The Administrative reform is best thought of as a cluster of initiatives with a general declared aim of improving accountability and effectiveness of the public sector.

But it is evident from the results of the completion of the cycle that administrative regulation and standardization of public functions and services has clearly dominated the rest of the initiatives.

The third cycle began in 2011 with the new set of initiatives that build on the results of the first and second cycle. It was accepted that the work has to be continued in most directions and that significant progress has been achieved mainly in administrative regulation of public functions and services and even there – with serious drawbacks.

The success in implementing performance management techniques has been humble and Russian commentators are generally sceptical on the results achieved so far. Some commentators consider that current performance management procedures may be considered "ritual" and "symbolic" (Malkov 2010). Others regard performance management an ineffective formalistic exercise (Khabaev 2010)

Commentators have reported issues of communication between reform design-teams. The formal division of the two clusters of reform initiatives led to the loss of coordination between expert groups involved in the implementation process and created problems of compatibility between reform streams (Barabashev and Klimenko 2010: 19). This has led to a reform programme that declares improved performance in service delivery as its strategic goal, but does not supply public sector managers with the means of cultivating a performance culture among individual officials (Verheijen and Dobrolyubova 2007: 210).

More broadly, it should be noted that, traditionally, performance management implies devolution of managerial discretion as one of its core elements (Batley and Larbi 2004). In the Russian case, however, declarations of principles of performance-based budgeting and performance-related pay somewhat paradoxically coexisted with vigorous efforts aimed at formalization and standardization of public sector functions:

> *The tension between home-grown solutions and the influence of international practice can be found in administrative reform, as well as other areas of political life. (Barabashev, 2007)*

It is as if performance management has been used as a ritual formula to indicate innovativeness and "western-ness" of the reformers' intentions, whereas their actual efforts

reveal the essential "eastern-ness" of the Russian public sector. Lip service was paid to ideas of decentralization and managerial autonomy to preserve the impression of liberalization, while real efforts were aimed at ensuring reliability and consistency in public service delivery by means of standardization and formalization[5].

The structure of government in historical perspective

It is worth giving here a brief outline of the structure of Russian government. Nominally, Russia is a Federation with 85 regions[6] (83 before 2014) (known collectively as *subjects of the Federation – субъекты федерации).* The executive consists of three tiers – municipal, regional and federal administrations. For purposes of this study it is important to consider the history of regional division and clarify the terminology. Writing about the Russian government in English necessarily entails a degree of simplification of existing Russian terminology. This section will try to introduce most important terms associated with the structure of Russian government.

I will use terms "regional administration" and "regional government" as "generic terms for the regional executive branch in Russia's regions" (Buckley et al. 2012: 13). The structure of regional administrations is the following:

> *The regional administration in most regions is comprised of three tiers, headed in all regions by a head of administration, colloquially called a "governor". Below that are the governor's deputies, colloquially called "vice governors." <...> From 1991-2006, regional chief executives in Russia's 21 ethnic republics were called Presidents (Buckley et al. 2012: 13)*

The structure of regional administrations vary, but it may be generally said to consist of ministries and the Office of the Governor (A*pparat Gubernatora*). Ministries are led by regional ministers and deputy ministers.

[5] An alternative interpretation would be that under the guise of "western" modern solutions, reformers were aiming at subordinating lower level bureaucrats by limiting scope for local discretion and "automatizing" service delivery.

[6] The number of regions gradually changed between 1991 and 2012: "Russia contained 89 federal subjects from 1991-2005. Between 2005 and 2008, the number was reduced to 83" (Buckley (2012: 13).

It is important to note that during the period covered by this study governors of Russian regions were appointed by the president rather than elected. Details of the procedure varied between 2007 and 2011 but the principle remained: governors were not elected but appointed. It is worth outlining the chronology of the institute of gubernatorial elections. Details may be found in (Buckley et al. 2012; Frye et al. 2011; Libman et al. 2012a):

> *From 1992-2004, Russia's governors were elected in a popular vote, but in 2004, then-President Vladimir Putin pushed through a reform cancelling direct gubernatorial elections and replacing them with a system of centralized appointment (Frye et al. 2011: 2)*

The main justification of this radical move was that there were too many amateurs and criminals among elected governors and that the country was torn apart by their local private interests:

> *Indeed, many justified the decision to cancel gubernatorial elections in 2004 as a way to purge criminals, political amateurs, and incompetents from the gubernatorial corpus (Frye et al. 2011: 3)*

Frye et al. (2011) cites Vladimir Putin's remark on the matter:

> *As Putin said in a July 2011 press conference, "I don't want to talk about [the cancellation of direct gubernatorial elections], but I will say: Back then…Everything seemed democratic and very good, but then pseudo-criminal elements began to rise up, fill their pockets with money, and manipulate both public consciousness and elected governors" (Frye et al. 2011: 3)*

Scholars differ in their assessment of the effects of this move. Frye et al. (2011) reviews two different camps: some argue that "elected officials are more likely to be of high quality because voters prefer high quality candidates" (Frye et al. 2011: 3), others hold that

> *…appointed officials may be of higher quality either because majoritarian failures result in political amateurs being elected… or because the social planners making appointments have long time horizons and seek the public rather than the private good (Frye et al. 2011: 3) .*

Libman et al. (2012a) outlines the chronology of Russian gubernatorial appointments:

> *Since the early 1990s the Russian regions have been ruled by well-entrenched politicians, spending many years (and even decades) in their region without any further option of advancement beyond the position of the governor and thus rather fitting the picture of a stationary bandit. In mid-2000s president Vladimir Putin abolished the public elections of the governors, replacing them by appointment by the centre. During the first years he had been rather cautious leaving most of the old governors in power. Over time, however, especially under Putin's successor in 2008–2012, Dmitriy Medvedev, the central re-appointment strategies became more aggressive. In the late 2000s, a new breed of regional governors came into existence: unlike their predecessors often recruited from regional elites, the new appointees usually came from high-ranked positions in the federal administration…*

It is important for the purposes of this study to keep in mind that the system of performance management operated between 2007 and 2012 and thus functioned under federally appointed governors.

Integral performance indicator

Some commentators have argued that starting from 2007 the central government used electoral results of the ruling party ("United Russia") in regional elections as an integral indicator of governors' performance[7]. If elections were confidently won the governor was considered able and a good candidate for promotions. This practice coincided with the introduction of the national system of performance indicators and, it is has been argued, foreshadowed it[8]. Election results became the one integral indicator of governors' performance. Political and administrative performance became intertwined and this led to manipulation of election results from bottom up. Electoral fraud became widespread because regional administrations tried to maximise the share of votes of the ruling party using their "administrative resource" (the power to influence local officials in election committees). It has been argued that large-scale manipulation of election results in 2011 was not a deliberate policy of the Kremlin, but an accumulated effect of uncoordinated manipulation of results by middle-tier civil servants who were all too eager to impress their superiors by exceptional performance.

[7] Yakovlev A.A. 21.02.12 "Ne tot signal podali gubernatoram" (The wrong signal to governors) http://slon.ru/russia/ne_tot_signal_podali_gubernatoram-749143.xhtml
[8] ibid

Alleged electoral fraud in 2011 parliamentary elections received a lot of public attention and even became subject of academic research (Enikolopov et al. 2012).

A note on terminology

In Russian managerial literature the English term "performance management" is commonly translated as «upravlenie po rezultatam» - «*управление по результатам*» (literary "management by results"). Such terminology persists, despite the fact that terms "results" and "performance" are not equivalent. Jakobson (2001) indicates that this terminology may "mislead":

> *The Russian translation of the term "performance management" is somewhat misleading …, the English term "performance" is roughly equivalent to the expression "quality of work" (kachestvo ispolneniya). Despite this, "performance" usually translates into "results" in Russian publications on the above topics (Jakobson 2001: 17).*

Mixing the concepts of "quality of work" and "performance" into "results" may lead to a somewhat narrow understanding of the subject. Current Western practice of performance management stems from such methods as Balanced Scorecard (BSC) (Kaplan and Norton 1992) and Total Quality Management (TQM) (Rogers 1996). The Balanced Scorecard approach, by its very title, implies that not only results are measured, but also other relevant aspects of organizational performance. TQM emphasizes the importance of the quality of the process, i.e. shifting the emphasis from the result itself to the *ability to* obtain high quality results *(capacity* to perform) (Dubnick 2005: 393). Thus, the Russian tradition of using the term "upravlenie po rezultatam" fails to reflect the intellectual legacy of the term "performance management". Performance management is an approach that is different from earlier doctrines such as "management by objectives" or "judgment by results". It builds upon these earlier managerial methods and incorporates the latest developments in organizational theory, including particular attention to capacity to perform, quality management, psychological and social dimensions of performance. The fact that the Russian managerial community adopted a term that fails to reflect

these aspects of performance may be explained by the absence of intellectual links with the Western managerial tradition. Thus, subtle semantic differences may indicate that ideas of performance management are not fully appreciated in Russia and are until today understood in the crude way that has been long discarded in the West as outdated. In Russia it is often understood that the core of performance management techniques is the use of formalized objectives, goals and targets, supplemented by incentives for achieving them. However this is only one part of performance management.

Although some researchers advocate distinguishing between a performance measure and a performance indicator (Jackson and Palmer, 1992, cited in Boyle, 1996: 6) I found this distinction of no particular use and have used the two terms interchangeably.

Performance measurement in the Russian public sector

Zhigalov, Pertsov and Chalaya (Zhigalov et al. 2009) identify two main categories of Performance Based Budgeting tools in Russian executive agencies: 1) goal-setting and planning tools that ensure the unity of policy 2) executive instruments that enable achievement of goals and objectives. In addition to PBB tools performance management instruments include strategic documents, performance assessment and monitoring instruments and incentivizing tools that are used to reward individual performance of civil servants (the last group of tools is only being formed in Russia.) All of these documents to a certain extent utilize performance indicators.

Table 1. Performance management framework in the Russian Federation

<table>
<tr><td>Federal strategic documents</td><td>Goal setting and planning tools</td><td>Administrative tools enabling achievement of goals and objectives</td><td>Performance Measurement Tools</td></tr>
<tr><td rowspan="2">• Concept paper of Socio-Economic Development of the Russian Federation until 2020 (КДР-2020)
• The main directions of activity of the Government of the Russian Federation (ОНДП)
• Development strategies (regional and sectorial)</td><td>Government programmes:
• Federal Targeted Programme (ФЦП)
• Federal Targeted Investment Programme (ФАИВ)
• Long-term targeted programmes (ДЦП)
• Departmental targeted programmes (ВЦП)</td><td rowspan="2">• Service quality standards
• Service needs assessments
• Service costs assessments
• Public service delivery agreements (goszadanie)
• Register of expenditure commitments (РРО)
• Justifications of budget appropriations (ОБАС)</td><td rowspan="2">• Annual report of the head of the highest regional executive body on achieved values of performance indicators. Decree 1199 (previously, Decree 825)[9]
• Annual report of the head of municipal administration on achieved values of performance indicators (Decree 607)[10]
• Reports on the results and main activities of administrators of budgetary funds (DROND, ДРОНД) (performance measurement section)</td></tr>
<tr><td>Reports on the results and main activities of administrators of budgetary funds (ДРОНД (DROND) (goal-setting section)</td></tr>
<tr><td></td><td colspan="2">Performance-based budgeting</td><td></td></tr>
<tr><td colspan="4">Performance management</td></tr>
</table>

[9] Presidential Decree of 28.06.2007 № 825 "On the evaluation of performance of the executive authorities of the Russian Federation".

[10] Presidential Decree of 28.04.2008 № 607 "On the evaluation of performance of local governments in urban districts and municipal areas".

Performance indicators used in this whole spectrum of documents are not harmonized. Some of them are used in several tools, others are unique to a particular programme or strategic document. This overlap (or the absence of one) presents certain difficulties when attempts are made to identify a suitable sample of performance indicators to be researched.

For the purpose of this study the sample includes performance indicators used in Performance measurement tools (Presidential Decrees 1199, 825, 607) and the section of DRONDs on performance measurement. To the extent that the same data are also used in government programmes and in policy documents and PBB tools, the study also covers these instruments (highlighted in grey in Table 1).

Performance measurement at the regional level

This study focuses on the practice of using performance indicators in Russian regional governments. Initiatives aimed at introducing elements of performance management into public sector operations may be traced back to 1998 (Blokhin 2011). This period may be characterized in terms of a shift in the focus of the federal government's attention away from managing political and economic crises towards managing and supporting steady economic growth and development. The federal government has turned to issues of efficiency and effectiveness and attempted to use performance management as means of rationalizing the public sector.

According to Blokhin (2011), at the outset, the federal government did not have any grand design of how the public sector should be structured. Reforms were implemented ad hoc to enable the civil service to absorb the unprecedented amount of money that flooded the federal budget. State incomes in the early 2000s grew as prices of oil climbed to record levels. As chronic deficits of public budgets were becoming a thing of the past, it was becoming more and more obvious that the civil service was unable to focus on strategic priorities; it was run in the mode of crisis management and had insufficient instruments or capacity to set and pursue clear objectives.

Among the activities undertaken by the federal government to overcome these challenges were performance management methods. First, they were tried out in federal ministries and agencies and then rolled out to regional governments.

This project explores the practice of performance management that emerged in Russian regional governments since the introduction of the first specially designed tool of performance management in 2004 (the DRONDs – Doklady of Resultatakh i Osnovnykh Napravleniyakh Deyatelnosti; rus. Results and Main Activities Reports) and was later extended by three separate sets of indicators (these were Presidential decrees 825[11], 607[12] and 1199[13]).

DRONDS were produced by regional administrations mainly for internal use and the set of indicators varied from region to region. The latter systems (colloquially referred to by shorthand "Decree 825 indicators" or "Decree 607 Indicators") were designed at the federal level. The focus of the quantitative part of this study is on Decree 825 indicators because only these indicators were systematically collected by the Ministry of Regional development and published in a single dataset. Decree 607 indicators (though most of them overlapped with decree 825) were reported by municipalities and were only published on websites of these municipalities. This made it almost impossible to compile them in a single dataset for analysis.

Decree 825 indicators cover all major areas of public services: Health, Education, Transport, Economy, Sport, Housing, Social care, Environment, Leisure facilities, Finances. The full list of indicators may be found in Appendix 6. They were published annually in May.

The total of 325 indicators were reported by 2011. 63 of them were included in the formula for grants. On the dynamics of these 63 indicators a league table was published. Top performing regions received grants (2bln rubbles ($67mln) to 20 top regions in 2007 and 10bln ($33mln) to 10 top performers in 2008-2011).

[11] Presidential Decree № 825 "On assessing the effectiveness of executive authorities of the Russian Federation.", 2007.
[12] Presidential Decree № 607 "On assessing the effectiveness of municipal authorities of the Russian Federation.", 2008
[13] Presidential Decree № 1199 "On assessing the effectiveness of executive authorities of the Russian Federation.", 2012

Table 2 lists the groups of indicators that were used in Decree 825. The largest number of indicators were related to Health care (80 out of 325), the smallest group was Security and Policing with only 5 indicators. This grouping was used in the decree that introduced the system.

Table 2. Groups of indicators collected according to Decree 825

Sphere	No. of indicators (% of total)	No. of indicators tied to grants (% of total)
Health care	80 (25%)	27 (43%)
Economic development	44 (14%)	6 (10%)
Vocational training and education	43 (13%)	0
Secondary Education	42 (13%)	16 (25%)
Communal services	40 (12%)	4 (6%)
Public administration	26 (8%)	7 (11%)
Housing	15 (5%)	3 (5%)
Energy efficiency	14 (4%)	0
Environment	10 (3%)	0
Roads	6 (2%)	0
Security and policing	5 (2%)	0
Total	325 (100%)	63 (100%)

To characterize the system of performance measurement created in accordance with degree 825, it is useful to use Christopher Hood's classification of performance measurement regimes into *target*, *ranking* and *intelligence* systems (Hood 2007b). *Target* systems aim at aggressively driving change by setting minimum standards or aspirational quantifiable goals and establishing sanctions and rewards. *Ranking* systems use league tables to identify winners and losers and apply appropriate measures. *Intelligence* systems aim at providing background information to facilitate learning without added pressure of formal targets and rankings.

The system combined *intelligence* and *ranking* types. A subset of 63 indicators was used by the Ministry of regional development to issue a league table of top performers. Coming first in the competition guaranteed no tangible "goodies" apart from rather small prize in the form of federal grants. In this respect the system differed from the British system of Comprehensive Performance Assessment (CPA) where good performance secured "freedoms and flexibilities" (Game 2006: 467):

> *[Excellent performers were] excused, for example, from producing certain statutory service plans for ministerial approval, less of their grant funding [was] "ring-fenced", and they [were] subjected to a "lighter touch" inspection regime (Game 2006: 467)*

The Russian system included no formal provisions for sanctions against poor performance. The data were used by the Presidential administration to evaluate governors and by governors to evaluate various departments of regional administration.

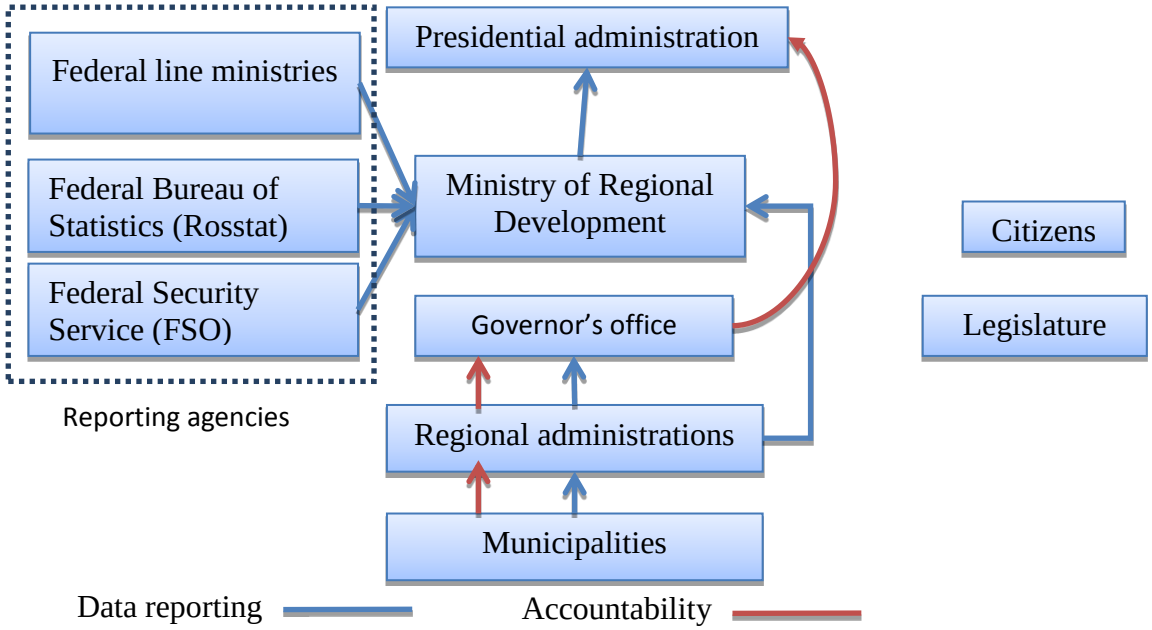
Data were published on official websites of regional administration and were available for the general public. However, the sheer volume of data made it almost impossible for laypersons to use it. In 2012 a data set covering the period between 2007 and 2011 was compiled and released by the Ministry of regional development.

Institutional context

Talbot (2008) used the term "performance regime" to speak about the institutional context of performance steering mechanisms. In the Russian case the following diagram would more accurately describe the network of institutions making up the performance regime (Figure 1). The legislative branch was not involved in the process of performance management. Respondents firmly stated that regional representatives never showed interest in the data generated under the Decree 825. This was chiefly because performance data from this system were not linked to the budget. Budgetary decisions were made separately through government programmes that had their own performance indicators, often dissimilar to those from Decree 825. Thus, the data stayed within the executive branch. Neither were the data actively used by

civic activists. This was, perhaps, because the data were presented in non-user-friendly way that made it hard for laypersons to use them.

Figure 1. Institutional context of performance reporting under Decree 825.



The type of transparency that this system established may be characterized as "bureaucratic" in the language of Christopher Hood (2007a): even though the data were published online and were nominally available for the public, they were presented in such a format that only specially trained civil servants could use them. Hood (2007a) introduces a simple 2x2 typology to distinguish between different types of transparency. Transparency may be either direct (observable by the people) or indirect (observable by experts or agents), it can apply to individuals or organisations. The relevant type for this study is indirect transparency applied to organisations (*bureaucratic* transparency in Hood's typology). The sheer amount of data was such that it made it nearly impossible for a layperson to make any sense of them. The system provided nominal transparency but in practice was inaccessible to the general public, so the data were generated by bureaucrats for bureaucrats.

Conclusion

In the period between 2007 and 2011 a system of performance measurement was implemented at the regional level in Russia. The system incorporated a list of over 300 indicators to monitor performance of regional administrations. On the basis of a subset of this performance data federal grants were distributed to top performers. The system included no formal sanctions for poor performance and operated largely in *intelligence* mode. The league table of top performers introduced an element of *ranking* into the system.

The introduction of a unitary national set of over 300 indicators overlapped with the prior initiative known as DRONDs (Reports on the results and main activities of administrators of budgetary funds (ДРОНД)). DRONDs were progress reports with a set of performance indicators compiled by departments of regional administrations in a bottom-up fashion. The set of indicators varied from region to region.

The national dataset collected in accordance with Decree 825 provided the opportunity to quantitatively study performance data generated by all 85 Russian regions.

CHAPTER 3. LITERATURE REVIEW

This section introduces major concepts that are later used in this study and reviews relevant literature on performance management and data manipulation. First, a brief introduction is made into the theory of bureaucracy and principal-agent theory; second, the role of performance management in wider New Public Management (NPM) movement is described; third, the literature on unintended consequences of measuring performance is surveyed; forth, the literature on data manipulation in both private and public contexts is reviewed; finally, the issues of data manipulation are placed in a broader context of public sector values to demonstrate the likelihood of perverse consequences of formal performance measurement in the Russian context.

Theory of bureaucracy

Tullock's rational choice theory of bureaucracy

> *The Politics of Bureaucracy provided the first-ever rational choice evaluation of the inner workings of a bureau, and it sets the scene for the more ambitious modelling and the statistical evaluations that would shortly follow (Tullock, 2005/1965, p.xvi)*

Gordon Tullock, an American economist and political scientist, published his book *The Politics of Bureaucracy* in 1965. In the introduction to a recent publication of the voluminous collection of his works Tullock is credited as one of the founders of the modern theory of bureaucracy. Tullock developed a simple model of the working of the hierarchy within bureaus. He used it to criticise the Weberian tradition of viewing bureaucrats as impartial servants of their sovereign (although it should be noted that Weber himself recognised that bureaucrats have interests of their own (Weber 1948)). Tullock's theory is a normative one, "[i]t is a theory that attempts to tell *how to make bureaucracies work*" (Tullock, 2005/1965, p.152, emphasis added).

In his model bureaucrats are reduced to mere "man units" (Tullock, 2005/1965, p.160). Such simplification allows him to abstract from individual personal qualities and motives that may guide bureaucrats in their action and concentrate on analysing superior-subordinate relationships in their pure form. He attempts to develop a general theory and discard "special

conditions" (Tullock, 2005/1965, p.161). He concentrates on superior-subordinate relationship in the hierarchy and the theoretical efficiency of bureaucratic structures in terms of organisational size, amount of time devoted to useful work and degree of control exercised by the sovereign. The general assumption that a typical bureau is a pyramidal structure with "a man at the apex" (Tullock, 2005/1965, p.162) is similar to Weber's.

Control in bureaucratic structures

> *The less that A has to control the activities of a given subordinate, the more likely are the activities of the subordinate to deviate from A's desires (Tullock, 2005/1965, p.154)*

Tullock views bureaucrats (and politicians) as self-interested individuals, who readily discard organisational objectives if it is conducive to attainment of their private benefits:

> *If the general atmosphere of his organization requires actions contrary to the attainment of the objectives of the organisation in order to secure promotion, the politician can hardly be expected to choose a course of action detrimental to his own advancement (Tullock, 2005/1965, p.44)*

This view would become the basis for future development of theories of bureaucracy, public choice and collective action as developed later by Niskanen, Downs, Olson and others. Tullock does not elaborate further in what respects bureaucrats' interests may be expected to differ from those of their superiors and the sovereign. This question was developed later by other theorists (primarily, Niskanen and Downs).

The problems raised by Tullock in his discussion of the working of bureaucratic structures are essentially those that modern systems of performance management attempt to address. Problems of conflicting objectives, supervision, control and coordination are intrinsic to bureaucracy and are as relevant today as ever before. Modern performance management doctrine is a new attempt to solve the same old problems.

Downs' personality types of bureaucrats

> *Downs's account of bureaucracy places motivational diversity at its core (Dunleavy, 1991, p.148)*

Downs, too, holds the view that bureaucrats maximise their utility. They optimize benefits net of costs: "every official acts at least partly in his own self-interest, and some officials are motivated solely by their own self-interest' (Downs, 1967, p. 83). And "every official is significantly motivated by his own self-interest, even when acting in a purely official capacity" (ibid. p. 262).

Limits to bureaucrats' rationality

An essential assumption in Downs' analysis of bureaus is that of bounded rationality. His theory differs in this aspect from neo-classical economic analysis:

> *"The bureaus and officials in our theory operate in a realistic world, not in the "perfectly informed" world of traditional economic theory. Therefore, even though we assume they make decisions rationally, there are limits upon their rationality. (Downs, 1967, p.75)*

Downs lists a number of limits to rationality of a decision making bureaucrat. These limits, in addition to their theoretical implications, have direct practical implications for the structure of performance management systems in the public sector:

1. *Each decisionmaker can devote only a limited amount of time to decisionmaking.*
2. *Each decisionmaker can mentally weigh and consider only a limited amount of information at one time.*
3. *The functions of most officials require them to become involved in more activities than they can consider simultaneously; hence they must normally focus their attention on only part of their major concerns, while the rest remain latent.*
4. *The amount of information initially available to every decision-maker about each problem is only a small fraction of all the information potentially available on the subject.*
5. *Additional information bearing on any particular problem can usually be procured, but the costs of procurement and utilization may rise rapidly as the amount of data increases.*
6. *Important aspects of many problems involve information that cannot be procured at all, especially concerning future events; hence many decisions must be made in the face of some ineradicable uncertainty. (Downs, 1967, p.75).*

Downs emphasises that not all officials have the same goals and that their goals may differ from that of the organisation:

> *In our theory, all large organizations are not teams, but coalitions. A team is a group of persons working together who have identical goals. A coalition is a group of persons working together who have some but not all goals in common. They need not give their common goals the same relative weight in their individual preference structures (Downs, 1967, p.76).*

Biased behaviour of bureaucrats

The central concept in Downs' analysis of bureaucracy is the concept of "biased" behaviour. According to his theory, a bureaucrat's behaviour in his official role inevitably has a certain bias (Downs, 1967, p.77). An official's *overall bias* measures the difference between the way he actually performs his roles in the bureau and the way he would perform them if his goals were identical with the formal goals of the organization.

Downs views a bureaucratic organisation as a hierarchy of principal-agent relationships where goals of principals and agents differ. As a result, "every organization usually has formal goals different from the actual goals of any of its individual members" (Downs, 1967, p.77).

According to Downs, "organisational goals" emerge as a result of a compromise between its members:

> *… "organizational goals" do not arise because the organization has a real personality independent of its members, or any "collective life" of its own. Rather they result from compromises among some or all individual members, who agree to adopt a formal set of goals not identical with the personal goals of any one of them. Perhaps there is no formal consensus about such "collective goals"; they may even be established by the fiat of the highest-ranking member of the hierarchy (Downs, 1967, p.77).*

It should be noted, that formal organisational goals within the context of a modern system of performance management are often set externally, or in consultation with external parties. Formal organisational goals may be set by politicians in limited consultation with members of the organisation. In such case, it is probable that "overall bias" of individual members of

organisation will be prone to increase. It is, thus, essential to envisage such potential problems in the process of designing systems of performance management.

Individual motives of bureaucrats

> *[The bureaucrat's] specific bias is always relative to some other particular official. It measures the difference between the way he actually performs his roles and the way he would perform them if his goals were identical with those of the other official concerned (usually his immediate superior or the topmost official in the bureau) (Downs, 1967, p.77).*

In addition to the overall bias, bureaucrats` behaviour may also be prone to specific bias that stems from differences in value systems and weighting of motives of individual bureaucrats. Downs identifies 2 types of motives that drive bureaucrats: self-interest motives and broader motives:

> *<u>Self-interest motives</u>*
> *Power - inside the bureau or outside it.*
> *Money income.*
> *Prestige.*
> *Convenience - minimizing personal effort.*
> *Security - defined as a 'low probability of future losses of power, money income, prestige or convenience'.*
>
> *<u>Broader motivations</u>*
> *Personal loyalty - to the immediate work-group, bureau as a whole, the wider government, or the nation.*
> *Identification with a specific programme of action or "mission-commitment"*
> *Pride in proficient performance of work.*
> *Desire to serve 'the public interest' - that is, what the official believes the bureau should be doing to carry out its social function. (Dunleavy,1991, p.148-9):*

The "utility functions" of bureaucrats are made up of both self-interest and altruistic goals (Downs, 1967, p.85).

Personality types of bureaucrats

It is important for the purposes of our further analysis to list here the bureaucratic personality types that Downs identifies: two "purely self-interested" and three "mixed-motive" types:

Purely self-interested officials are motivated almost entirely by goals that benefit themselves rather than their bureaus or society as a whole…

***Climbers** consider power, income, and prestige as nearly all-important in their value structures.*

***Conservers** consider convenience and security as nearly all-important. In contrast to climbers, conservers seek merely to retain the amount of power, income, and prestige they already have, rather than to maximize them.*

Mixed-motive officials have goals that combine self-interest and altruistic loyalty to larger values. The main difference among the three types of mixed-motive officials is the breadth of the larger values to which they are loyal.

***Zealots** are loyal to relatively narrow policies or concepts, such as the development of nuclear submarines. They seek power both for its own sake and to effect the policies to which they are loyal. We shall call these their sacred policies.*

***Advocates** are loyal to a broader set of functions or to a broader organization than zealots. They also seek power because they want to have a significant influence upon policies and actions concerning those functions or organisations.*

***Statesmen** are loyal to society as a whole, and they desire to obtain the power necessary to have a significant influence upon national policies and actions. They are altruistic to an important degree because their loyalty is to the "general welfare" as they see it. Therefore, statesmen closely resemble the theoretical bureaucrats of public administration textbooks (Downs, 1967, p.88).*

Problems that arise from the divergence of individual and organisational goals are not specific to public sector organisations, they exist in the private sector as well. However, in profit-making organisations profit may serve as an objective measure of performance and, thus, provides a way of detecting and limiting biases among employees. "But the equivalent limits in bureaus are far more obscure and uncertain" (Downs, 1967, p.78).

A system of performance management that seeks to improve performance of public sector organisations need to take into account these considerations regarding individual motives and goals of officials. For, if a performance management system is to influence civil servants behaviour, it needs to be able to predict it, and "in order to predict what officials will do, we must know their goals" (Downs, 1967, p.82).

Bureaucratic rationality

Knowing the underlying motives of civil servants is indispensable in designing a system of managing employee performance. Knowing "what levers to pull" is crucial in attempting to manage performance. Thus, individual motives of public sector officials should be discussed in greater detail. Each motive may be linked to a particular tool within the system of rewards and penalties. Thus, the consequences or redesigning systems of performance evaluation may be predicted if such motives are known. Weber, for example, held the view that secure monetary reward and clear career prospects are integral to effective working of a bureaucratic hierarchy:

> *According to experience, the relative optimum for the success and maintenance of a strict mechanization of the bureaucratic apparatus is offered by a secured money salary connected with the opportunity of a career that is not dependent upon mere accident and arbitrariness (Weber, 1948/1922, 208).*

The relative stability of employment and security of money income are characteristic features of the public sector. Public sector employees may attach great value to steadiness and predictability of career progress. This means that such elements of performance management systems as regular performance reviews, formal assessments of employee's performance and other forms of tests are highly important.

There are two extremes to the system of employee performance evaluation. One is the ideal Weberian type: impartial, impersonal and rational:

> *[The bureaucratic apparatus] develops the more perfectly the more the bureaucracy is 'dehumanized,' the more completely it succeeds in eliminating from official business love, hatred, and all purely personal, irrational, and emotional elements which escape calculation. This is the specific nature of bureaucracy and it is appraised as its special virtue (Weber, 1948/1922, 216).*

This logic prescribes such systems of performance evaluation that are unaffected by either the superior's or the subordinate's personalities. They should be objective and meritocratic. They require the results of one's work to be objectively measured. The idea of quantitative performance measurement of results and performance indicators is in line with this logic.

The other extreme is complete arbitrariness in evaluating one's performance. In this case, one's career prospects are left to their superior's discretion. This is in conflict with the principal of bureaucratic rationality and impersonality and is characteristic of less developed systems of administration, like the administration of notables in feudal states.

Of course, in practice such extremes do not exist in public sector organisations. There is necessarily a balance between objectivity and subjectivity in evaluating performance of personnel. Because of the hierarchical nature of bureaucratic organisations, one's performance is often evaluated by one's superior: "…a significant proportion of evaluative decisions regarding bureau personnel must be based on subjective personal judgements. The man most effectively placed to make such judgments (from the point of view of the bureau) is each official's immediate superior" (Downs, 1967, p.81). In the following sections evidence will be presented showing that in the Russian case often arbitrariness prevails and civil servants often are promoted not based on formal criteria but based on their superiors' discretion.

Performance management as an element of NPM doctrine

Performance management as an element of New Public Management

The label "performance management" encompasses a diverse range of managerial methods used to define goals, create plans, allocate managerial discretion, measure outcomes and reward performance (Heinrich 2002; Ittner and Larcker 2001; Kravchuk and Schack 1996; Otley 1999).

It is useful at this stage to distinguish between performance measurement, performance reporting and performance management. These terms have often been used interchangeably, but a sharper distinction may be useful in understanding both intended and unintended consequences of performance measurement. The following definitions are suggested by Radnor and Barnes (2007: 393):

> (1) *Performance measurement* is a quantitative or qualitative value of the input, output, outcome or level of activity of an event or process.

(2) *Performance reporting* is providing an account, and often some analysis, of the level of input, activity, output or outcome of an event or process usually against some form of target.

(3) *Performance management* is action, based on performance measures and reporting, aimed at improvements in behaviour, motivation and processes and promotes innovation (Radnor and Barnes 2007: 393).

Performance management is one of the building blocks of New Public Management (NPM) (Hood 1991). (Batley and Larbi 2004) conducted a survey of publications on NPM and identified the pool of elements that may be considered part of NPM doctrine (Table 3). Within these typologies two principles lie in the domain of performance management: a) to create clear standards and performance indicators for enhanced accountability and b) to shift the focus of attention from processes to output controls[14] (Gray and Jenkins 1995; Hyndman and Eden 2000, 2001; Lapsley 1999; ter Bogt 2003). Between these two functions there is often assumed to be some synergy. However, some researchers have questioned that and pointed out contradictions that arise between these two functions in practice (Bouckaert and Halachmi 1994; Bouckaert and Peters 2002; Halachmi 2002a, 2002b). These contradictions have been termed "accountability paradox" (Dubnick 2005).

Dubnick (2005) studied the logical connection between accountability and performance, and concluded that the existence of synergy between these functions does not necessarily exist, whereas contradictions may be predicted theoretically and have been recorded by empirical research. He concludes that the assumption of a positive link between accountability and performance is often merely rhetorical, not empirical.

[14] Official documents that frame performance management reforms in Russia assume these functions to be complementary See, e.g., the Concept paper for the administrative reform in the Russian Federation 2006-2010. («Концепция Административной реформы в Российской Федерации в 2006-2010 годах»). The implementation of performance management, according to this document, should allow for 1) "*more effective control* over the execution of decisions of federal and regional agencies", and 2) "*increased effectiveness* in the use of resources by federal and regional agencies". (Degree of the Government of the Russian Federation #1789-r, as of 25 October 2005).

Hood, 1991	OECD, 1995	Ferlie et al., 1996	Borins, 1997	Osborne and Gaebler, 1992
Hands-on professional management	Decentralized management environment to replace centralized, hierarchical structures	Decentralization: organizational unbundling; new forms of corporate Governance: move to board of directors model	Increased managerial autonomy; reduce central controls	Decentralized government authority – flexible, less layered forms of organization
Shift to disaggregation of units into quasi-contractual or quasi-market forms		Split between strategic core and large operational periphery		Catalytic government: 'steering not rowing'
Shift to greater competition and mixed provision, contracting relationship in the public sector; opening up provider roles to competition	Creation of competitive environments within and among public sector organizations; flexibility to explore cost-effective alternatives to direct provision and regulation	Elaborate and developed quasi-markets as mechanism for allocating resources within the public sector	Maintaining receptiveness to competition and private sector participation	Competition within public services: may be intra-public or with a variety of alternative providers
Stress on private sector styles of management practice	Closer focus on quality, feedback from clients and other interests groups	A split between public funding and independent service provision	Creating synergy between the public and private sectors	Driven by mission/goals rather than rules and regulations
Greater emphasis on output controls	Increasing focus on results – efficiency, effectiveness and quality	Stress on provider responsiveness to consumers; major concern with service quality	Providing high – quality services that citizens value: service users as customers	Customer driven; empowering citizens
Explicit standards and measures of performance	More focus on efficiency with establishment of productivity targets	More transparent methods to review performance	Demanding, measuring and rewarding both organizational and individual performance	Result-oriented government: emphasizing outputs not inputs
Stress on greater discipline and parsimony in resources use; reworking budgets to be transparent in accounting terms	Strengthening the ability of the centre to guide and respond better to change and diverse interests at least cost	Strong concern with value-for-money and efficiency gains	Provision of human and technological resources that managers need to meet their performance targets	Enterprising government: earning money not just spending it
		Downsizing		Leveraging change through market rather than bureaucratic mechanisms
		Deregulation of the labour market		Anticipatory government: prevention of problems rather than treatment

Performance management

Table 3. Performance management among other elements of New Public Management. Source: (Batley and Larbi 2004: 42) selection added.

The assumption of a positive link between accountability and performance is derived from the principal-agent model of relations between politicians/citizens and bureaucrats. It is assumed that greater accountability, clearer objectives and more transparent budgeting provides principals with better tools of control and, thus, leads to greater performance. This is in line with Tullock's theory of bureaucracy. This model assumes that bureaucrats (agents) need to be controlled and

supervised to make sure they work in principal's interests. This model also assumes that principals want greater control and are able to utilize it effectively.

However, control is not free. It is costly on both sides. Acts of reporting involved in control mechanisms require agents' energy, time and money that could be used for other activities, for example, to work directly on the managerial task. Principals need to process and digest information. In addition, control and oversight also limit the "freedom of action" of the agents. If agents are more competent than principals, their performance may be undermined by too much principal's attention. The costs of excessive accountability may outweigh benefits, creating a situation in which control hinders performance instead of stimulating it.

On the other hand, too little control is also dangerous. Single-minded pursuit of the 3Es that characterizes New Public Management may have detrimental effect on democratic principles of accountability of government to the public (Hood 1991). Maximum results may be achieved by abandoning accountability completely (Piotrowski and Rosenbloom 2002) but this is not in the society's interests.

Russian policy documents on performance management traditionally list control and efficiency as two main functions of performance management. This simplification endures despite numerous scholarly publications that attempt to separate the two functions and practical evidence suggesting that in reality often one function dominates.

Why measure? Managerial and political benefits of PM

> *…the function does not explain the existence, but the existence does serve the function. (Bryant, 1991:23)*

In accordance with "Durkheim's injunction to consider separately the cause of something and the function it fulfils" (Bryant, 1991:23), it is necessary to separate functions (or benefits) of performance measurement systems in the Russian regional government and causes of their introduction. Functions (or benefits), in turn, may be intended and unintended, formal and informal (declared and non-declared). One should not completely disregard officially declared

functions of performance indicators, for '[w]hile the functionality of a given institution is never a complete explanation of that institution, … arguments about functionality … constitute an aspect of a proper explanation' (Wright, 1983:15)

Performance management may provide two types of benefits: managerial and political (de Bruijn 2002; Frank 2008; Kloot and Martin 2000; Propper and Wilson 2003).

Managerial benefits are associated with the improvement in the organization's work process, results and overall performance. Behn (2003) identified eight managerial purposes of performance measurement:

> *As part of their overall management strategy, public managers can use performance measures to evaluate, control, budget, motivate, promote, celebrate, learn, and improve. Unfortunately, no single performance measure is appropriate for all eight purposes. (Behn 2003: 586)*

These eight managerial purposes are summarized in the table below:

Table 4. Eight managerial functions of performance measurement. Source: Behn (2003)

Managerial function	**Meaning**
Evaluate	How well is my public agency performing?
Control	How can I ensure that my subordinates are doing the right thing?
Budget	On what programmes, people, or projects should my agency spend the public's money?
Motivate	How can I motivate line staff, middle managers, non-profit and for-profit collaborators, stakeholders, and citizens to do the things necessary to improve performance?
Promote	How can I convince political superiors, legislators, stakeholders, journalists, and citizens that my agency is doing a good job?
Celebrate	What accomplishments are worthy of the important organizational ritual of celebrating success?
Learn	Why is what working or not working?
Improve	What exactly should who do differently to improve performance?

Frank (2008: 430) identifies four managerial "purposes" of implementing performance management: clear objectives and targets help employees focus on important aspects of agency's performance ("communicative purpose" (Kaplan 2001; Rangan 2004)); measurement and publication of performance data allows the public to see what is being achieved for their money ("transparency purpose"); performance measures may be used to learn and improve ("learning

purpose"); performance measures may be used to reward individual performance ("appraisal purpose").

Bird et al. (2005: 4) identified three "broad but diverse" aims of PM: "finding out 'what works, identifying the functional competence of practitioners or institutions, and public accountability".

Performance indicators can facilitate communication between public sector organizations and communication within organization. Performance targets provide a way of transferring lofty mission statements into measurable outputs (Hyndman and Eden 2000; Kaplan 2001; Rangan 2004). So, even though performance data are not always directly used to improve managerial decisions or alter budgetary allocations, they can help to identify and constructively discuss potential problems (Melkers 2006).

Political benefits are the benefits that lie beyond the sphere of management (Kloot and Martin 2000; Meier and O'Toole 2006; Propper and Wilson 2003). Hood and Dixon (2010) identified three groups of political benefits: 1) ideological, 2) symbolic and 3) direct electoral benefits. Ideological benefits may include the achievement of a desired transformation, e.g. the advance of ideology which the politician shares. Symbolic benefits include ''the ability to communicate . . . that government is being run in a rational, efficient and results-oriented manner and that bureaucrats are being held accountable for their performance'' and therefore improving media image and chances of re-election (Hood and Dixon 2010: 282). This helps to build a favourable image of the power and increase the chances of re-election (Moynihan 2008: 68). Direct electoral benefits "involve improved re-election chances for incumbents from public management policies that convey benefits to key voters or shape the policy agenda in electorally favourable ways" (Hood and Dixon 2010: 282). Another type of political benefits may also be discerned: the effects of performance management reforms may include a shift in the balance of power within ruling elites, e.g. centralization or decentralization of power (by establishing

channels of accountability and formally mapping superior-subordinate relationships within bureaucratic elites).

Performance indicators may assume particular political significance in competitive democracies (Johnsen 2005:14): "PIs may effectively function as carriers of information, functioning as 'prices' in political markets, in much the same way as prices do in input and product markets". Political parties may be highly interested in performance reports, if they can use performance figures strategically in political competition.

Hood and Jackson (1994) provide a historical overview of successive public administration doctrines over the past two hundred years. They conclude that, in choosing between two managerial doctrines, technical superiority plays less important role than political passions of the moment. Political fashions change as each following generation of politicians attempts to portrait their ideas as the most advanced, while degrading the ideas of previous generations by accusing them of being outdated and erroneous. The pendulum of public policy needs a strong nudge to swing in the opposite direction and it often swings too far before it is swung back again.

The global popularity of performance management is an example of such "fashion"(Kelly 2002) There are conflicting pieces of evidence regarding managerial usefulness of performance management in the public sector, but, nonetheless, the desire to implement performance indicators is not withering away. Modern incarnation of PM as an element of NPM is only one in a series of similar initiatives that appeared on the agenda of governments throughout the 20th century: "the idea of enhancing accountability through scientific approaches to management is a cultural constant just like the appetite for reform" (Kelly 2002: 377). Kelly (2002) argues that this is due to symbolic benefits of "political rhetoric". Performance management "became the public administration orthodoxy of the period rather than a set of testable propositions" (Kelly 2002: 377). Politicians believe that citizens want to hear about the results of the government and that all efforts are applied to improve these results. Performance indicators bring semblance of

rationality in politics, and, thus, serve to create the impression of "effective management"[15]. Civil servants may also be interested in perpetuating a rational image of their organisations by implementing policies that are generally regarded as rational:

> *'By structuring itself along lines that are generally regarded as reasonable, fair, efficient, rational, modern and so on, an organization can win the understanding of its environment' (Brunsson (1989: 4)cited in Pollitt (2001: 941)).*

Performance data use

> *Performance indicators are a valuable managerial tool if properly used; if not, managerial time and cash resources could he wasted, and, more seriously, managerial action could be distorted (Likierman, 1993, p.15)*

Why are performance data actively used in some contexts but not in others? Recent implementation studies focus on explaining how performance information is used to generate both symbolic and instrumental benefits for managers of public sector organisations or for external stakeholders (Poister and Steib, 1999; Pollitt, 2005; Modell, 2004; Moynihan, 2005; Taylor, 2007, 2011; Torres and Yetano, 2011).

The place of implementation in the wider context of reform

Implementation is only one of the stages in the life-cycle of a reform. The process of implementation is shaped by a number of external factors. Pollitt and Bouckaert (2000) schematically depict the process of policy implementation (Figure 2 - N). Reforms are launched in response to socio-economic challenges (block A). There are three elements in this block: (B) – global economic conditions; (C) – socio-demographic shifts and (D) national socio-economic policy at the macro level.

[15]Examples of the use of performance indicators for this purpose may be found in both Russian and Western practice. Two examples may illustrate this. In USA President Clinton commented on signing the Government Performance Results Act in 1993: "The law simply requires that we chart a course for every endeavor that we take the people's money for, see how well we are progressing, tell the public how well we are doing, stop the things that don't work, and never stop improving the things that we think are worth investing in." In Russia prime-minister Vladimir Putin in 2011 referred to performance indicators during an extended TV interview. He was asked whether career decisions concerning regional governors were made based on electoral results on the ruling party "United Russia" in the governor's region. He denied this by referring to "objective, balanced, purely industrial" indicators that are used to evaluate a governor's performance. He claimed that these indicators (introduced in 2007 by Presidential decree #825) are the true criteria for governors' promotions or demotions. Similar instances were observed during the fieldwork stage of this research project: during public hearings and discussions of regional strategic documents civil servants used performance indicators to justify budget decisions on rational grounds.

Figure 2. The structure of the discourses surrounding implantation of a government policy (Pollitt and Bouckaert 2000)

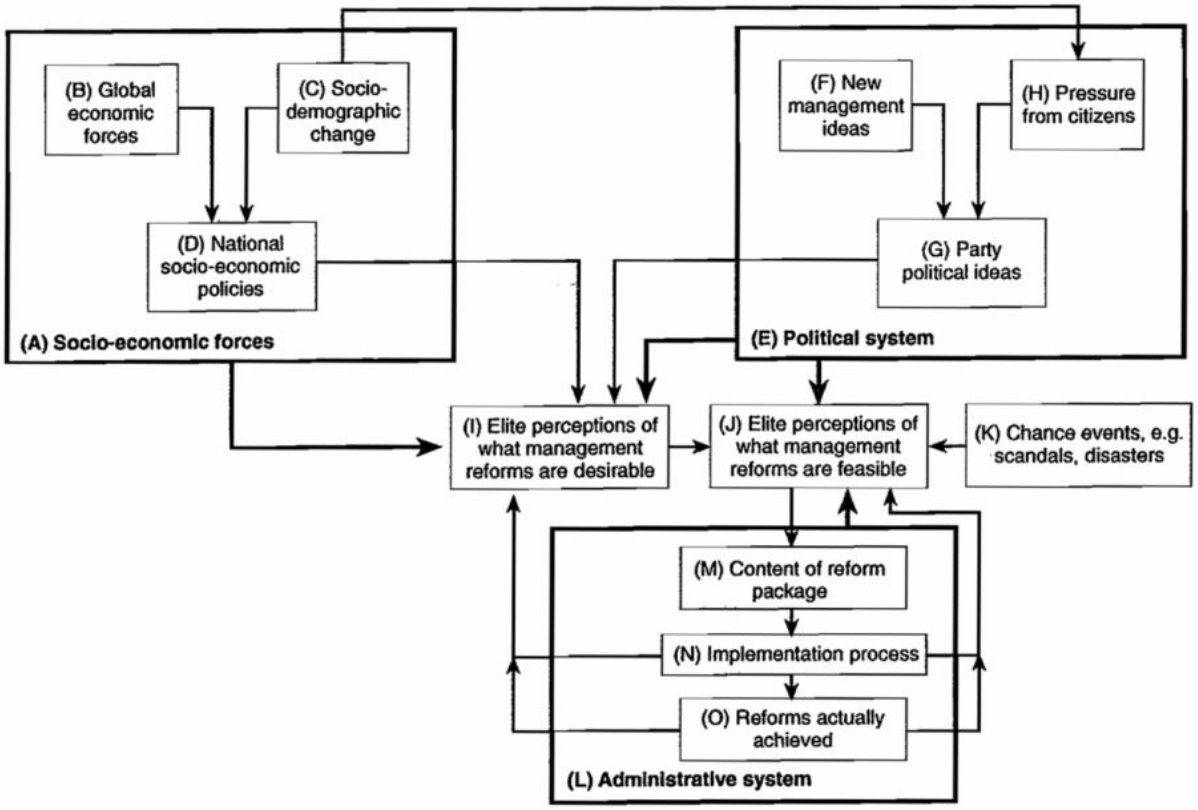


FIG. 2.1. A model of public management reform

Socio-economic challenges require political actions. The model shows how these actions are brought about in western democracies. Governments come under pressure from citizens to design solutions to pressing problems of the day (H). This creates a demand for new managerial solutions (F). These new managerial solutions are filtered through party ideologies and crystallise in government programmes and policies (G).

The model underlines the role of "elites" in the process: political, economic and social elites determine what policies are adopted. Taking into consideration socio-economic conditions (A) and the political agenda (E) the elites develop a particular desired reform package (I).

Then this package goes into implementation phase and comes under the influence of bureaucrats. During the process of implementation initial proposals are shaped and moulded to

fit interests of the administrative system (L). Other factors (scandals, resource constraints, changing political agendas (K)) also shape what is actually achieved (O). The results of the implementation process are then used by the elites to change their expectations and form a new reform package (I) and reiterate the process (J to O).

Pollitt and Bouckaert (2000) highlight the importance of elite perceptions of what reforms are feasible and desirable. In forming a policy it is not the objective state of affairs that is important but the perception of this state by various elites (political, economic, bureaucratic) (J) and (I). In other words, the formulation of a reform programme is a discursive process with many subject positions and stakeholders.

For the Russian case the model should be amended to reflect the relatively lower role of political competition and pressure from citizens. Nonetheless, in general, the process proceeds through the same stages.

Internal demand vs external pressure to implement PM systems

Agency may implement performance measures in response to external requirements or due to some internal demand. In the former case it may happen that performance indicators and managerial decisions become decoupled (Radnor 2008). "Ownership" of PM is likely to be lost and performance reporting degenerates into a "box-ticking" exercise. Radnor (2008) provides a useful illustration how lack of ownership, "box-ticking" and lack of understanding of target baseline combine to turn managers into mere administrators of a hollow performance measurement exercise:

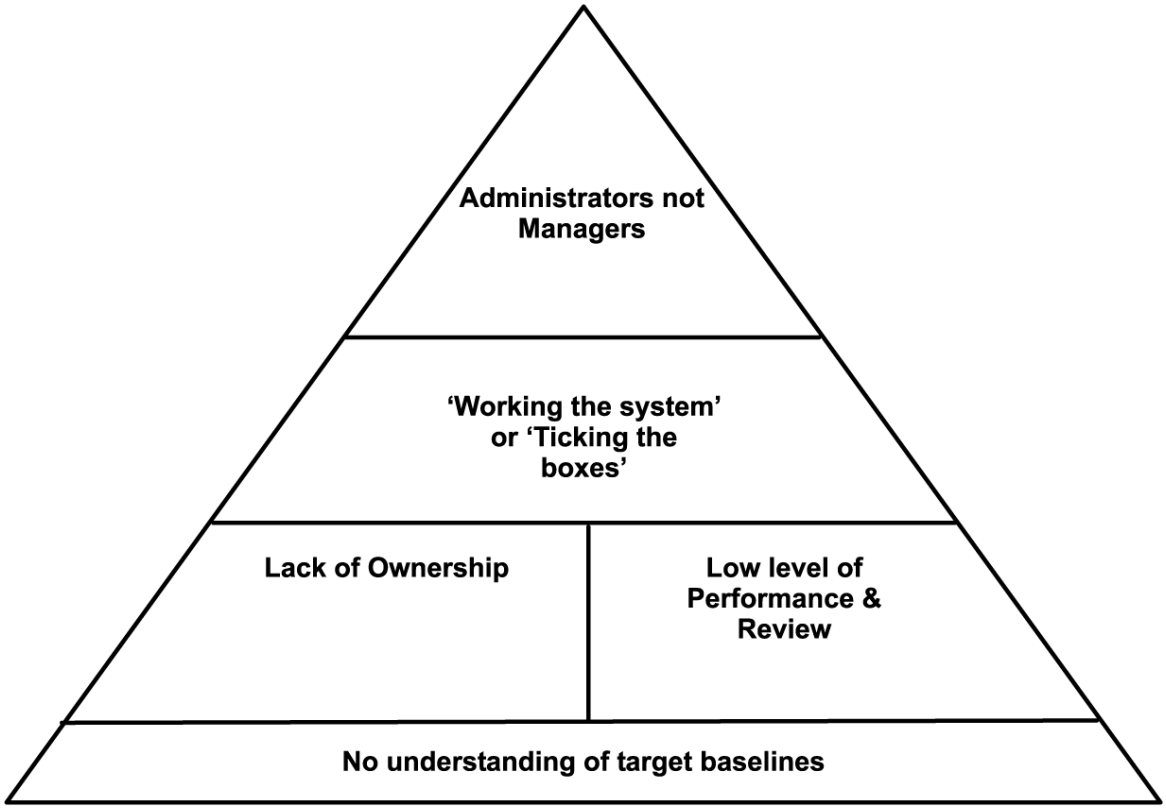


Figure 3. Combination of factors reducing public sector managers to mere administrators of a PM regime. Source: Radnor (2008)

The mere fact that an agency produces performance data is not a sufficient condition for a real performance management culture to emerge (Taylor, 2007). A study of government organisations running KPI demonstrated that "the sheer existence of KPIs in public agencies' annual reports and budget papers does not automatically lead to their effective use by the authorities for making decisions" (Taylor, 2007:341). She found that "[Australian public sector] agencies use PIs more for meeting external reporting requirements than for achieving internal improvements" (Taylor, 2011:860).

Even when performance management systems are introduced in response to internal demand the rationale for their implementation may be symbolic rather then managerial: Torres (2011) in a study of 7 biggest Spanish cities found that "the symbolic or institutional image value of taking decisions under rational decision-making models is the primary driver of implementation processes" (Torres, 2011:1081). Melkers and Willoughby (2005), Taylor (2007, 2011) and Torres (2011) agree that as long as performance indicators are implemented as a

response to external accountability requirements they tend to be ignored for purposes of internal decision making. If performance measures were introduced by higher authorities to strengthen control, they were not used for management decisions, and vice versa, if managers introduced their own measures, regulators ignored them (Torres et al. (2011).

Performance reporting as an administrative procedure may continue despite the lack of interest in published data. Poister and Steib (1999) found that in the United States only "40 percent or fewer [of municipalities with more than 25 000 population] make any kind of meaningful use of performance measures in their management and decision processes" (Poister and Steib, 1999:332). The survey showed that the requirement to introduce performance indicators did not come from the federal government or state governments: "the overwhelming motivation to use performance measures… [stemmed] from a desire to make better decisions and to maintain accountability to citizens and local elected officials" (Poister and Steib, 1999:333). Even though it seemed that performance measurement was introduced to meet demands of local stakeholders, only 20% of respondents indicated that performance reports were actually communicated to such stakeholders.

Even where performance management has become part of common managerial wisdom, performance data may receive little attention. Pollitt (2005) in a cross national study of 4 European countries concludes that, despite the fact that "in north western Europe, performance measurement has become almost universal", its use remains an activity "conducted by and for managers": "politicians do not take much interest, and neither do citizens" (Pollitt, 2005:41).

Among the barriers for effective performance data use there is also a lack of capacity of local public servants. Boyne et al. (2002) studied performance data reporting in municipalities of Wales. They concluded that, due to the lack of skills and experience of staff, "documents such as performance plans [made] little contribution to the accountability of public organizations"(Boyne et al. 2002: 691).

Over more than 30 years of active efforts at implementing performance management a number of lessons have crystalized. Early on it was pointed out that, if performance management is to become a fully fledged managerial practice, it should necessarily be accompanied by significant changes in managerial culture: managers should be free to manage (Likierman, 1993). Likierman based his conclusions on data from over 500 interviews of "middle- and senior-grade public service managers" He identified 20 lessons of successful implementation of performance management. Among Likierman's many lessons one is particularly relevant to the Russian case. Likierman recognises that the implementation of performance indicators cannot be successful, if it is not accompanied by "fiscal devolution" and decentralisation of managerial decision making. One of the interviewees observed that

> *without other managerial changes, for example financial devolution, the performance indicator culture will not flourish… PIs would merely bob along the top of the organisation making the occasional appearance on the agenda of management teams (Likierman, 1993, p.20).*

Other lessons emphasise the importance of feedback and involvement of the staff and other stakeholders into the process of designing and setting performance targets. This emphasis on feedback, negotiation and involvement is shared by many modern implementation studies and culminates in the recent concept of public governance and network governance (S. P. Osborne 2009).

Difficulties in implementation of performance management are not uniquely Russian. In developed Western economies things also did not always go right. Two years after Likierman's study Meekings (1995, p.5) observed that "partial coverage, poor implementation and unrealized expectations [were] still the norm" among agencies implementing PM. He pointed that "insufficient attention [was] paid to the process of implementation" which he called "the key to unlocking the real potential of performance measurement" (Meekings, 1995, p.6). Meekings calls for greater focus on the process of developing measures, rather than on measures themselves:

Organizations may have the most elegant measures in the world, but if people don't actually want to use them they are worthless (Meekings, 1995, p.6).

Assessing quantifiable benefits of PM

Many of the studies mentioned above use case studies or literature reviews to arrive at their conclusions; others focus on a small sample of public sector organisations or a local performance management initiative of relatively small scale. There is another group of studies that attempt to quantitatively evaluate the effects of two major nation-wide PM initiatives in the USA: GPRA (Government Performance Reporting Act) and PART (Programme Assessment Rating Tool). These initiatives have been introduced at the federal level in the United States to assess the performance of federal agencies (GPRA), and federal government programmes (PART). GPRA was put into practice in 1993, PART - in 2002. The former was the initiative of the Congress, the latter – that of the president. Despite the fact that considerable time has passed since the implementation of these two initiatives (20 and 10 years respectively), the debate about their usefulness continues (Yang 2011).

Some authors argue that there is a serious contradiction between the two instruments (Gueorguieva et al. 2009), others question their cost-effectiveness (Dubnick 2005). Researchers disagree in their estimates of long-term effects of these performance management tools. Some point to the fact that civil servants gradually learn to use performance information effectively (Newcomer 2007), while others find no visible impact of reforms (Gilmour 2008). At the same time, the sheer amount of performance data generated under GPRA is in itself a limitation for a productive use of this information (Breul 2007: 328). The value of the data is not obvious, as not only citizens, but also members of the Congress show little interest in it (Breul 2007: 328).

PART was presented by the administration of President George W. Bush as a "neutral, objective tool" for "a thorough, systematic and transparent "evaluation of government programmes" (Yang 2011). However, its neutrality and objectivity have been put into question by some researchers (Dull 2006; Gilmour and Lewis 2006; Greitens and Joaquin 2010; Radin 2008). Consistency of evaluation has also been questioned (Gueorguieva et al. 2009). It has been

demonstrated that some programmes with low PART-ratings received additional funding, while funding for programmes with high ratings decreased (Gilmour and Lewis 2006; Radin 2008). This has been attributed to an ideological bias of the presidential administration. Under republican presidents a positive correlation between PART ratings and funding decisions has only been observed for "traditionally democratic programmes". "Traditionally republican programmes", on the contrary, showed no such link. The authors theorized that this could be the result of deliberate "insulation" of ideologically favourable programmes from the effect of negative PART evaluations. For democratic programmes, on the contrary, PART-ratings could be used for supporting essentially political decisions with "objective evidence".

Unprecedented growth in the amount of performance data produced by public sector agencies and improvements in statistical software facilitated quantitative studies of performance management (Andrews et al. 2008; Jackson 2011). Frank (2008) gives an extensive overview of recent quantitative studies. Publication of performance data allowed researches to assess the link between performance and other variables such as the measure of decentralization (Andrews et al. 2009) and target-setting (Boyne and Chen 2007) However, despite the use of sophisticated techniques of quantitative analysis, results remain controversial (Boyne 1998; Boyne et al. 2005). Results are often sensitive to model specification, and can easily be "massaged" to fit any preconceived conclusion using equally plausible assumptions. Further complications for quantitative analysis emerge from the presence of external factors. It is often challenging if not impossible to separate the influence of one particular reform on the operations of the public sector. Reforms rarely come alone, most commonly a set of reforms is implemented simultaneously accompanied by numerous external influences. Thus, for example, even though performance data in the UK's public sector organizations demonstrates a positive trend, it is not simple to find what role in this trend is played by the introduction of performance targets (Boyne and Chen 2007; Clarkson et al. 2009).

Unintended consequences of performance measurement

> *…objective performance measures may have certain dysfunctional results, particularly when important parts of an official's tasks are immeasurable (Downs, 1967, p.80).*

Performance measurement systems have been advocated in the public sector on the grounds of their potential benefit in terms of efficiency and effectiveness (Smith, 1995, p.277). It has, however, been documented that the introduction of performance measurement systems may have adverse unintended consequences on the organisation that is subjected to monitoring (de Bruijn 2002; Hood and Peters 2004; Hood 2007a; Thiel and Leeuw 2002; Vakkuri and Meklin 2006). Ridgway (1956) discussed dysfunctional effects of quantitative performance measurement over fifty years ago. More recently, a number of reasons of unintended consequences were summarized by Smith (1995).

The basic idea behind formal performance measurement is "the notion of managerial cybernetics", that is – the view of the managerial process as a cybernetic cycle:

> *Organisational objectives are identified. Performance indicators are developed to reflect these objectives. Targets are set in terms of the performance indicators. Management then chooses action and effort intended to achieve the targets. Progress towards targets is monitored using the PIs and – if there is a divergence from targets – new targets are set and appropriate remedial action is taken, so the process continues (Smith, 1995, p.280)*

In this framework managers of the organisation are seen as agents under supervision of their principal - the body that designs and operates the system of performance indicators. Smith (1995, p.280) uses the principal-agent theory to argue that the main purpose of any system of performance indicators (PIs) is for the principal to establish organisational control over the agent through the process of getting feedback and making decisions based on that feedback. Unintended consequences emerge because targets and indicators are used to "control a complex network of self-controlling human beings", therefore "the system that one is trying to control is intelligent" and "can anticipate the actions of the controller… and take action to frustrate [the controller's] wishes" (Smith, 1995, p.280).

Principal-agent theory suggests that the ability of the principal to make the agent do what the principal wants depends largely on the congruence of the principal's and the agent's objectives. Any divergence in the objectives of the two parties undermines the ability of the principal to secure maximum effort of the agent (Smith, 1995, p.283).

Another important general consideration here is the acknowledgement of the fact that no system of performance measurement is completely neutral. The very act of measurement has the potential to influence the process that is being measured. This has long been acknowledged in the literature on performance management in the corporate sector: "[w]hat gets measured gets attention, particularly when rewards are tied to the measures" (Eccles, 1991). This increased attention may or may not be indented by the designers of the PM system.

Based on this understanding of the functioning of performance measurement systems Smith (1995, p.183) identifies eight possible adverse unintended consequences of the introduction of a system of performance indicators*: tunnel vision, suboptimization, myopia, measure fixation, misrepresentation, misinterpretation, gaming and ossification*. *Gaming* is further analysed by Hood (2006) and Bohte and Meier (2000). The former uses the Soviet experience in central planning to highlight three most significant types of gaming*: ratchet effect, threshold effect and goal displacement*.

There is no one established classification of unintended consequences of PM in the literature and often similar actions are given different names by the commentators. The fact that such unintended consequences do occur is, however, widely accepted (Bird, 2004, p.1).

A brief summary of these unintended effects is given in the table below:

Table 5. Unintended consequences of measuring performance. Sources: (Smith, 1995; Bevan and Hood, 2006).

	Definition given by Smith (1995) and others
tunnel vision	Smith (1995, p.284) defines *tunnel vision* as "an emphasis by management on phenomena that are quantified in the performance measurement scheme, at the expense of unquantified aspects of performance"
suboptimization	"Suboptimization is the pursuit of narrow local objectives by managers at the expense of the objectives of the organization as a whole". (Smith, 1995, p.286).

myopia	"…PIs can induce managerial myopia, the pursuit of short term targets at the expense of legitimate long term objectives". (Smith, 1995, p.288)
measure fixation	"…**measure fixation** can be defined as an emphasis on measures of success rather than the underlying objective" (Smith, 1995, p.290)
misrepresentation	"Misrepresentation is the deliberate manipulation of data so that reported behaviour differs from actual behaviour" (Smith, 1995, p.292)
misinterpretation	… bounded rationality might cause the controller systematically to misinterpret [the data], and to send the wrong policy signals to the agent (Smith, 1995, p.294).
ossification	"…ossification: organizational paralysis brought about by an excessively rigid system of performance evaluation" (Smith, 1995, p.299).
Gaming:	Gaming is the term that refers to behaviour patterns that agents adopt in order to maximise their benefits under the existing performance measurement regime.
ratchet effect	*'…a wise director fulfils the plan 105 per cent, but never 125 per cent ' (Nove, 1958 , p. 4 cited in Bevan and Hood, 2006, p.521).* Managers who know that their next year's targets are based on last year's performance have the perverse incentive not to exceed the target (Bevan and Hood, 2006, p.521).
threshold effect	This effect refers to the situation in which successful organisations, whose performance prior to the introduction of a target was higher than the target, have the perverse incentive to reduce their performance and align it with the majority of average performers. (Bevan and Hood, 2006, p.521).
goal displacement	Same as *tunnel vision* (alternative term offered by Bevan and Hood, 2006, 521)

Of particular interest for our project is the type of unintended consequences labelled by Smith as *misrepresentation*. Smith defines misrepresentation as "the deliberate manipulation of data so that reported behaviour differs from actual behaviour" (Smith, 1995, p.292) The possibility of misrepresentation exists because of the information asymmetry between the agent and the principal. The agent is better informed and in some cases may have the discretion over the way in which an event is recorded and reported. The agent can, therefore, choose such a way of recording an event that will portray his performance or efforts most favourably. Smith (1993, p.148) gives an example of a doctor who has the opportunity to maximise his apparent workload by allocating a patient into one or another of diagnosis-related groups.

Misrepresentation can also take the form of direct fraud or lying. A high profile example of such "performance enhancement tactic" is "the exaggeration of body counts in Vietnam" (Bohte and Meier, 2000, p.175). Such examples are abundant in modern press, particularly in cases

related to military operations or street protests. Each side attempts to portray itself in a more favourable light by collecting and reporting "objective" statistical data. A possible strategy of mediating this type of unintended consequences is external audit of the reported information. It is, however, costly and time-consuming (Smith, 1993, p.148).

Costs of enforcing compliance

Unintended consequences may undermine effectiveness of performance measurement regimes and hinder policy implementation. Enforcement may be used to ensure that agents comply with formal requirements of a policy. According to principal-agent theory the cost of such enforcement and the scale of damage done to the policy by unintended consequences depends on the actors' propensity to exploit the system to their advantage. The greater the discrepancy between goals of agents and goals of the organisation the harder and more costly it is to enforce compliance and reduce the scale of undesirable practices. In the end, any formal requirement and limitation may be circumvented, if people are sufficiently motivated to do so. Commenting on the efforts of financial regulators to limit the extent of irresponsible practices in financial markets Yavlinsky (2011:129) remarked, that "[h]uman imagination, whipped up by the prospects of easy money, will always find a way to circumvent formal constraints". Similarly, in the public sector, if a system of rewards and sanctions is so designed that it creates motivation for individuals to engage in gaming or other forms of unintended practices, such practices would emerge.

Costs of enforcing compliance are in direct proportion to the individuals' propensity to abuse the system. "[E]nforcement of compliance with the rules… is ineffective unless… agents discharge them without duress" (Yavlinsky, 2011:38). Agents may be forced to comply with the rules by coercion. However, the cost of such coercion becomes prohibitive if the scale of exploitation grows sufficiently high. If agents are willing to observe the "rules of the game" voluntarily, the cost of ensuring compliance will fall dramatically:

> *If, say, nine economic agents out of ten stick to the rules not because of daily … coercion but because they consider it fair and rewarding, violations of these norms can be prevented or punished by authorized institutions in accordance with majority sentiment, with coercion applied only when strictly necessary. As a result, general compliance with the rules of the game … is achieved at minimal cost (Yavlinsky, 2011:38).*

According to public choice theory, in the context of the public sector performance management, the propensity of individual civil servants to comply with the rules would depends, first, on their personal motives and attitudes (such as those described by Downs (1967) in his typology of bureaucrats: some bureaucrats possess commendable moral principles and their system of goals converges with that of the organisation, others are driven mostly by self-interested motives and disregard organisational goals); and, second, by the ability of policy implementators to ensure "buy-in" of bureaucrats. Many implementation studies emphasize the importance of "ownership" of the policy by staff, managers and other stakeholders (Audit commission, 1999; Meekings, 1995; Likierman, 1993; Commonwealth Secretariat, 1994a, b, 1995, Kim, 2002). Such ownership may be ensured through wider participation, engagement and consultations: "When you consult people, you put things on the agenda. When you involve them, they make up their own agenda" (Audit commission, 1999:48).

In real life there is never a perfect fit between organisational and individual objectives, moreover, systems of performance management, instead of bridging the gap, may increase it. This happens when a performance management regime creates perverse incentives and stimulates organisational cheating. One of many ways public sector managers may cheat is through deliberate manipulation of performance figures to show their organisations in a more favourable light.

Data manipulation in public and private sectors

In discussing incentives for data manipulation it is useful to use Christopher Hood's classification of performance measurement regimes into *target*, *ranking* and *intelligence* systems (Hood 2007b). *Target* systems aim at aggressively driving change by setting minimum standards

or aspirational quantifiable goals and establishing sanctions and rewards. *Ranking* systems use league tables to identify winners and losers and apply appropriate measures. *Intelligence* systems aim at providing background information to facilitate learning without added pressure of formal targets and rankings. Most evidence on data manipulation comes from research in *target* systems (this literature links back to Soviet Union economic debates about ratchet and threshold effects in planning (Hood 2006)). Comparatively little is known about incentives for data manipulation in *ranking* and, particularly, *intelligence* systems (Hood 2007b).

A large body of literature is devoted to methods of detecting data manipulation. Two major streams are relevant for the present study: 1) data manipulation in the public sector within performance measurement regimes; 2) manipulation of financial data in the private sector.

Data manipulation in the public sector: gaming or cheating?

The fact that formal performance measurement may lead to various unintended consequences has long been recognized in the literature (Ridgway 1956). It is, however, not easy to distinguish between harmless effects and potentially harmful practices. It is hard to draw the line between a legitimate exercise of managerial discretion in recording outputs and illegitimate organizational cheating. A number of typologies of perverse effects of performance measurement have been proposed. Many papers in the field build on Smith's (1995) classification of unintended consequences to construct typologies of organizational gaming.

Radnor (2008) suggested a 2x2 typology: the axes are "impact of gaming" and "level of gaming within the organisation." The four types "move from muddled to massaging to manoeuvring then manipulated" (Radnor 2008: 324).

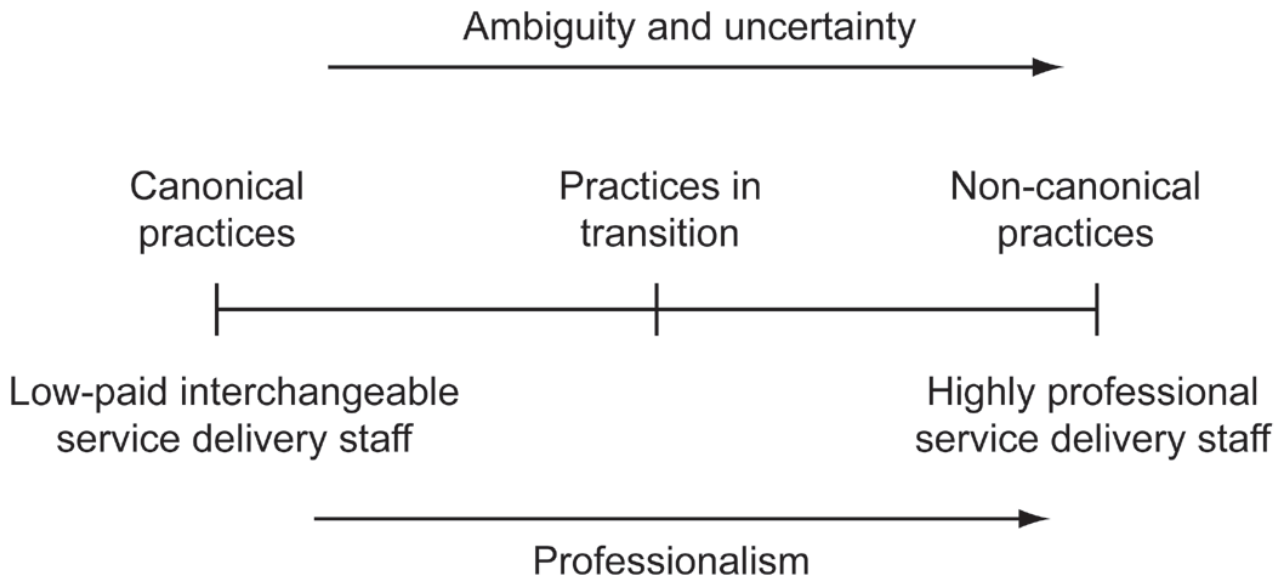
	Impact of gaming	
Level of gaming within organization	Internal	External
Low	Muddled	Massaging
High	Manoeuvring	Manipulated

Table 6. Typology of organisational gaming. Source: (Radnor 2008)

Here "Manipulated" represents the situation where the pressure to meet targets is so great that it distorts service provision significantly. An example is the infamous case of patients being left waiting in ambulances in order to meet the Accident and Emergency department 4 hour waiting target (the clock did not start ticking until the patient was inside the hospital). On the other hand, "Muddled" represents the case where the external impact of the performance reporting regime is small and the scope of manipulation is not significant. When the data are used only internally and targets do not create distortionary pressure on service delivery. An example of this type is a museum that counted all people entering the building as visitors despite some of them being contractors or staff. This is likely when performance measures are ambiguous.

Pidd (2005) stresses that performance measurement was at the heart of UK government's efforts to improve public services under Tony Blair. Not surprisingly, a large number of research articles have since investigated the British experience. Pidd (2005) argues that perverse effects of performance measurement reflect the "Goodhar's Law": "any observed statistical regularity will tend to collapse once pressure is placed upon it for control purposes" (Pidd, 2005:486). Pidd (2005) uses two different typologies to explain the emergence of unintended consequences of quantitative performance measurement. The first is Noordegraaf and Abma (2003)'s typology of public management practices (Figure 4). This typology groups public management practices into three types: *canonical*, *practices in transition* and *non-canonical* practices. *Canonical* practices are characterized by low ambiguity and are more suitable for quantification, measurement and standardization. These practices may be performed by low-paid interchangeable service delivery staff. *Non-canonical* practices are characterized by high ambiguity and are performed by highly professional service delivery staff. These practices are hard to measure because significant aspects cannot be quantified and standardized. Pidd (2005) argues that unintended consequences of performance measurement are likely to emerge when attempts are made to subject *non-canonical* public sector practices to quantitative measures and targets.

Figure 4. Ambiguity, uncertainty and non-canonical practices. Source: Pidd (2005)

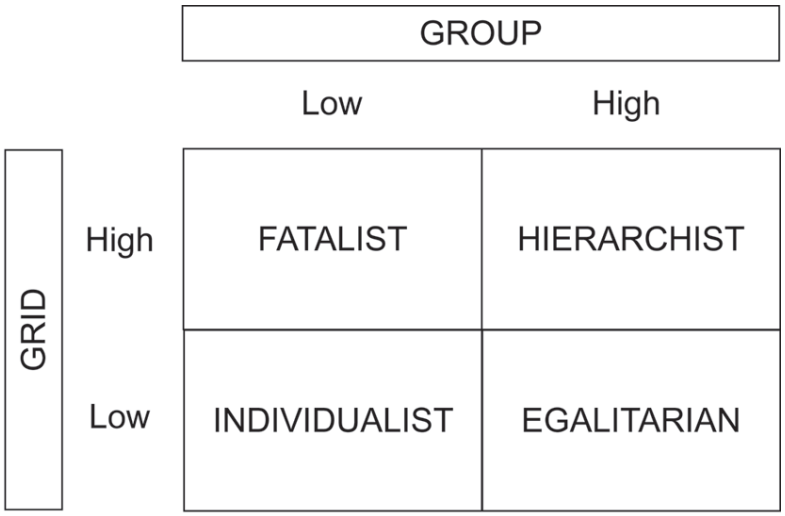


Another typology used by Pidd (2005) is Hood's (1999) typology of organizational cultures. Hood's typology is a simplified version of the typology proposed by Douglas (1982). This is a typology of dominant patterns of social interactions in social actors' world. These patterns are categorized on two dimensions:

> *(1) Grid. This indicates the degree to which people's actions are governed by externally imposed rules and conventions. If grid is high, there is little scope for individuals to negotiate what they do and how they do it.*
>
> *(2) Group. This indicates the degree to which actions are governed by group choice that is the social collectivity to which someone is committed. (Pidd 2005: 490)*

Again, a 2x2 typology is suggested (Figure 5)

Figure 5. Simplified grid: group theory. Source: Pidd, 2005



This grid: group typology produces four types of "social worlds":

> *(1) Hierarchist. This is a highly prescribed world in which external rules and group norms tightly define what is permissible. Thus, personal security and predictability are highly valued.*
>
> *(2) Individualist. This lies on the same diagonal as hierarchist and is therefore a world in which individuals have great freedom of choice but may have little security.*
>
> *(3) Fatalist. In this world, the individual is placed in tightly prescribed roles with little or no autonomy and no obvious way to cooperate with others in the group.*
>
> *(4) Egalitarian. In this world there are no clear rules defining the roles of members, but there is a clear distinction between those who are part of the group and those who are not. Active participation in group activities is expected. (Pidd 2005: 490)*

Pidd (2005) then argues that combining this typology with Noordegraaf and Abma's typology may be useful in predicting perverse consequences of formal performance measurement. When performance measures are imposed "from on high" on an organization that is characterized by fatalist social relations, then it is likely that operations of these organization would be distorted. External requirements would override weak internal standards of the organization. If the staff is low-paid and used to dealing with canonical practices and low uncertainty levels, it is likely that external pressure to achieve quantitative target would overrule the internal drive for proper service provision. Pidd (2005) argues that many observed harmful effects of performance measurement are due to the fact that performance measures are set

externally for organisations that lack internal capacity to deal with them responsibly. In hierarchist organisations (where conformity to internal rules is high) it is unlikely that external performance measures would significantly distort behaviour. However, if fatalist behaviour dominates, external targets are likely to produce perverse distortions such as those listed by Smith (1995).

Financial data manipulation in the private sector

Meanwhile, a large body of literature have developed around manipulations of financial data (known as "earnings management") (See, for example, (Beatty et al. 2002; Bennett and Bradbury 2010; Burgstahler and Dichev 1997; Dechow et al. 1995; Degeorge et al. 1999; Eames and Kim 2012; Jansen et al. 2012; J. J. Jones 1991; K. L. Jones et al. 2008). Researchers focused on accounting practices that made it possible for managers to adjust earnings to meet expectations of analysts and shareholders. Sophisticated methods of modelling have been proposed to evaluate the scope of potential manipulations.

Findings of researchers in this area are relevant to the study of performance data manipulations. Beneath the surface of dissimilarities between financial and performance data there lie underlying patterns of human behaviour that are quite similar in both cases. So, although, the phenomena that is being measured may be different (cash flows vs length or roads or number of beds in hospitals), the incentives to manipulate measurements and associated interpretations of the figures are quite similar. In both cases the resulting measures reflect not only something that may be called "objective reality" (the dynamics of cash flow or of provision of some municipal services), but also the "edifice" or super-structure of human interventions linked to the power struggle (or interplay of expectations and blame avoidance strategies) between management and shareholders, or between different levels of bureaucratic hierarchies. This "superstructure" acts as a prism distorting the image of objective underlying reality. Configuration and inner workings of this superstructure are quite similar in both cases, so that insights about the nature of the one may help explain and predict behaviour of the other.

Since this body of literature is hardly ever referenced in debates on public sector performance measurement, it is worth reviewing it is some detail. A number of studies focused on strategies of "earnings management". Jones (1991) found that managers deliberately used their discretion over "income-decreasing accruals" to reduce their companies reported income during periods of "import relief investigations" (special procedures administered by the United States International Trade Commission):

> *Explicit use of accounting numbers in import relief regulation provides incentives for managers to manage earnings in order to increase the likelihood of obtaining import relief and/or increase the amount of relief granted" (Jones, 1991, p.193).*
>
> *Earnings management arises from the game of information disclosure that executives and outsiders must play. Investors base their decisions on information received from analysts—usually indirectly. Say, through a broker— and through published earnings announcements. To bolster investor interest, executives manage earnings, despite the real earnings sacrifice. (Degeorge et al. 1999: 3)*

Jones (1991, p.194) argues that in the case of receiving import relief there is an asymmetry of incentives: a defuse group of losers has low incentives to monitor managers, whereas a concentrated group of winners has incentives to manage earnings. In other earnings management studies, he argues, stakeholders may have more balanced incentives to monitor dishonest practices:

> *Income relief is a wealth transfer from a group of diffuse losers (consumers) to a group of concentrated winners (all other contracting parties of domestic producers receiving import relief). I argue that consumers do not monitor earnings management as effectively as losers examined in other studies because the loss to each consumer is smaller, and their interests more diverse, than for the contracting parties examined in these studies. Regulators have less incentive to adjust for managers' earnings manipulations since their ultimate payoff for such adjustment is less direct than in other situations previously studied (e.g., union contract negotiations). (Jones, 1991, p.194)*

In another study Dechow (1995) evaluated a number of statistic models that test for earnings management and found that none of them could reliably detect "earnings management of economically plausible magnitudes" (Dechow, 1995, p. 223). Large samples of several

hundred companies are required to achieve robust conclusions. In any case, it is not possible to infer from statistical results that a particular firm engaged in dishonest earnings management.

In an influential article Degeorge et al. (1999) introduced three theoretically predicted thresholds of earnings management: to show positive profit, sustain recent performance and meet analysts expectations. They base their predictions on the assumption that managers deliberately manipulate earnings in their self-interest or in the interest of shareholders:

> *Earnings provide important information for investment decisions. Thus executives—who are monitored by investors, directors, customers, and suppliers — acting in self- interest and at times for shareholders, have strong incentives to manage earnings. We introduce behavioural thresholds for earnings management. A model shows how thresholds induce specific types of earnings management.* ***Empirical explorations identify earnings management to exceed each of three thresholds: report positive profits, sustain recent performance, and meet analysts' expectations*** *((Degeorge et al. 1999: 1) emphasis added).*

Degeorge et al. (1999) find that positive profit threshold dominates and that firms suspected of boosting earnings show poorer performance in the long run. This may be contrasted with the findings of the present study. According to my findings, sustaining recent performance is the dominant motive of data manipulation in the public sector ("prudent" data manipulation appear to dominate). Civil servants prefer to maintain status quo, whereas private sector managers prefer positive growth.

The main indicator of suspicious behaviour are the so called "kinks" in distribution of frequencies of earnings increases. Unusually high frequency of small increases and unusually low frequency of small decreases indicate potential manipulations (see figures below). Such "kinks" violate statistically predicted distributions (such as Benford's law (Durtschi et al. 2004))

Burgstahler and Dichev (1997) find evidence for the hypothesis that "incentives to avoid earnings decreases become stronger with the length of the previous run of earnings increase" and, thus, irregularities in the distribution of earnings are sharper among firms that demonstrated earning increases in a number of previous years (Figure 6). Each year stakes get higher. With each consecutive year of positive profits the pressure to demonstrate positive growth builds up. If earnings were manipulated once, it becomes imperative to manipulate them again and again to

maintain the positive image. Thus, a gap between real and reported performance may appear. The burden of potential blame gets heavier and heavier each time the indicator is manipulated.

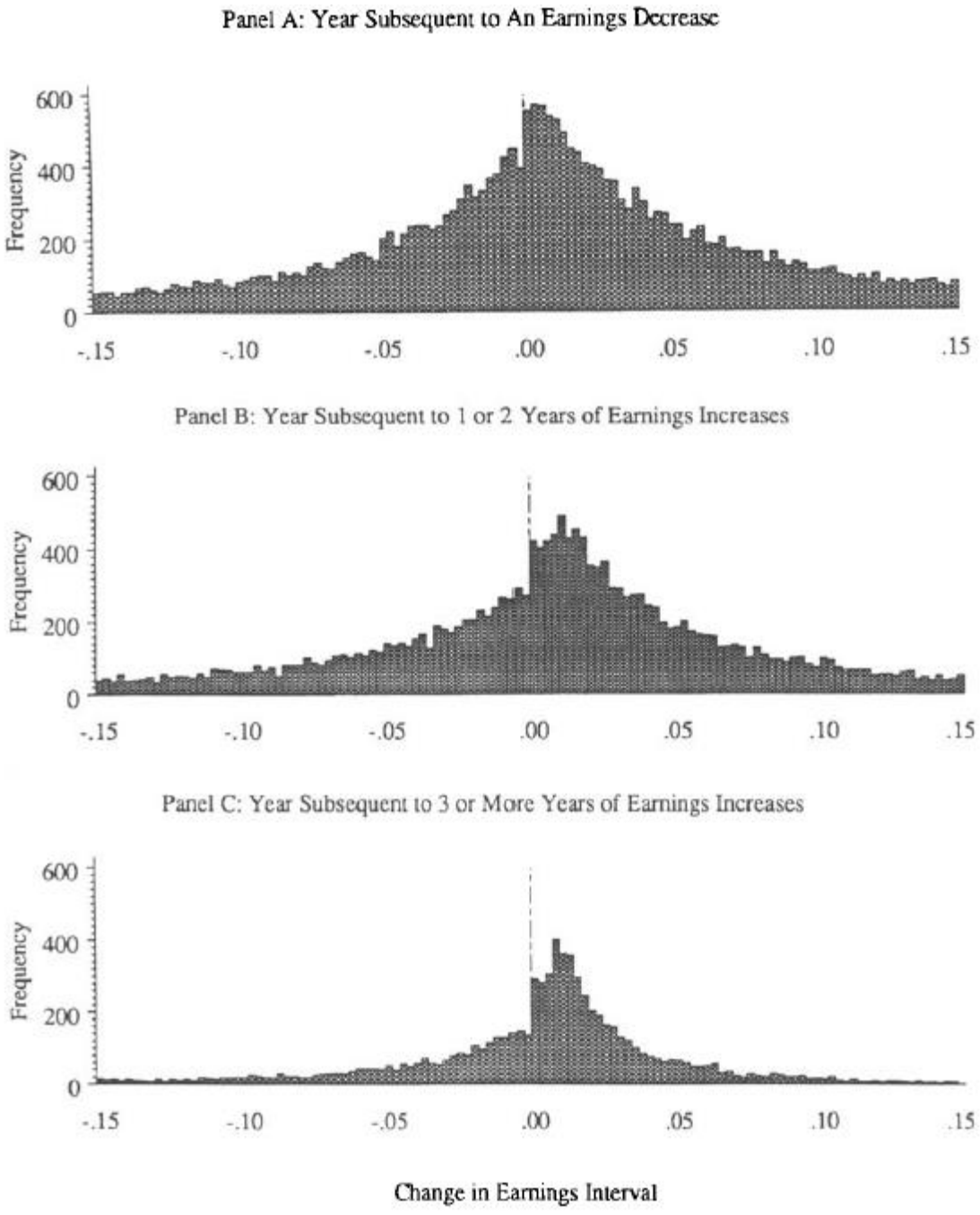


Figure 6. "Three empirical distributions of changes in earnings scaled by market value categorized according to the pattern of preceding earnings changes for the firm.

Panel A: the distribution for the years immediately following an earnings decrease; Panel B: the distribution for the years following one or two years of earnings increases; Panel C: the distribution for the years following three or more years of earnings increases". Source: Burgstahler (1997: 106)

(Beatty et al. 2002) found that publicly owned banks are more likely to report small earning increases and tend to report longer strings of consecutive earnings increases than private banks

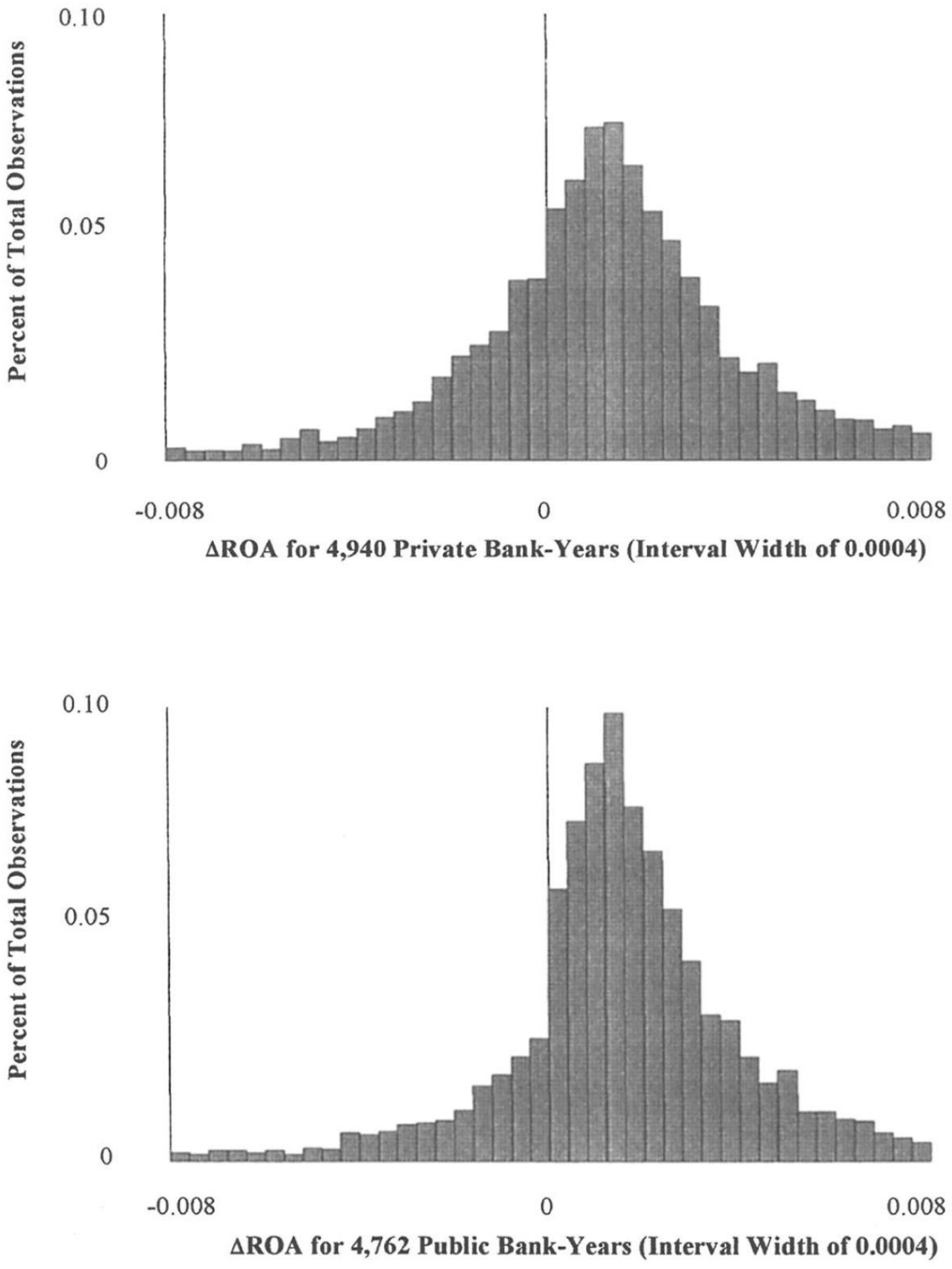


Figure 7. Earnings of public and private banks. Source (Beatty et al, 2002: 560)

They attribute this to the fact that "public banks' shareholders [are] more likely than private banks' shareholders to rely on simple earnings-based heuristics in evaluating firm performance" (Beatty et al. 2002: 547). Privately owned banks are likely to rely on more elaborate indicators

and place larger weight on non-financial performance. (Beatty et al. 2002) found supporting evidence for the earnings management hypothesis.

Blame avoidance and trust in performance data in the public sector

The public sector is generally believed to have a strong blame avoidance culture. One of the major contributors to the study of performance management theory and practice Christopher Hood explores the consequences of the mixing performance targets with blame avoidance mentality of the public sector (see, for example, (Hood 2002, 2007b, 2007a, 2011). Hood emphasizes that in their attempts to "manage by numbers" the British government had to face many of the challenges of the Soviet system of economic planning. He focused on practices of "gaming" in NHS hospitals and also used anecdotal evidence from other spheres. Other authors explore the implications of risk-averse mentality in the public sector on systems of performance management/measurement (Charbonneau and Bellavance 2012; Rozin and Royzman 2001; Rubenstein et al. 2003).

Charbonneau (2012) studied performance data from Quebec's Municipal management indicators regime and found that lower performing agencies are more likely to use the provision of justifications to explain their performance. The authors found that blame avoidance behaviour was pronounced "even in a regime with few incentives, no consequences linked to performance, and limited transparency to citizens" (Charbonneau and Bellavance 2012).(p.319). This suggests that even when performance indicators are being implemented as an intelligence tool they suffer from the consequences of negativity bias. Authors argue that negativity bias is prevalent in the public sector and results in deliberate usage of blame avoidance techniques. Naturally this raised the issues of trust. If performance figures are adjusted to show an agency in the more favourable light, trustworthiness of such data may be questionable. Yang and Holzer (2006) argue that performance measurement in the public sector can potentially improve citizen trust in government, but often fails to do so because of lack of citizen engagement with government

agencies. In addition to lack of general trust in government performance data themselves may be untrustworthy.

This is particularly relevant for countries where the government dominates the economy. Historically in the Soviet Union management by numbers was an integral part of large bureaucratic hierarchies of state. There is a number of exemplary cases from the Soviet experience that are traditionally used to illustrate problems of data manipulation: the so called "cotton case" and the "Ryazan case". The cotton case (or the "Uzbek case") was a high profile instance of manipulation of economic indicators in the 1980s in what is today Uzbekistan (K.E. 2009). Indicators of amount of cotton produced were systematically inflated to demonstrate stable growth of productivity. When the manipulations were discovered a number of high-level party functionaries were convicted. The Ryazan case (or "Ryzan miracle") is another Soviet example of unintended consequences of performance targets (Agarev 2005). Collective farms were given target amounts of meat to produce. In order to meet the targets local party leaders sanctioned culling of milk-producing cattle and confiscations of private cattle. Short-term performance was boosted at the expense of long-term sustainable performance. Production sharply fell the next year and those party functionaries who were promoted for outstanding performance were subsequently convicted of criminal offence.

It has been noted that in modern day China similar reasons make national statistics unreliable. Regional heads are able to manipulate performance figures to secure promotions. It is unclear to what extent this is the case, but it raises serious concerns. Chinese economy is one of the powerhouses of modern global economy and Chinese economic growth figures influence economic decisions worldwide. Koch-Weser (2013: 4) found "anecdotal, statistical, and legal evidence [suggesting] that many local statistics cannot be trusted". Because of both objective

(lack of reliable methods of data collection) and subjective (desire of civil servants to report better performance) factors Chinese economic statistics should be taken with a pinch of salt[16].

Other spheres of data manipulation

Instances of discovered data manipulation practices are abundant in modern Russian press. The table below provides a sample of such news items from a popular news website:

Table 7. News articles on the subject of data manipulation.

Title	Date	Summary
Lenta.ru: Интернет и СМИ: Интернет: Что так сердце, что так сердце раскрутило	21.08.2013	Alleged data manipulation during a survey for the "symbol of Russia". The governor of Chechnya – R.Kadyrov allegedly funded fraudulent votes Chechen nominee – the mosk of Grozny.
Lenta.ru: Россия: Общество: ФСБ обнаружила фальсификацию результатов ЕГЭ в Кабардино-Балкарии	08.06.2013	Manipulation of school test results during Unified National Exam (ЕГЭ).
Lenta.ru: Россия: Преступность: Петербургскую полицию уличили в искажении отчетности	06.04.2013	Data manipulation in crime statistics.
Lenta.ru: Россия: Общество: Хлопонин потерял 110 тысяч детей на Северном Кавказе	14.03.2013	Alleged manipulation of the number of orphans in Northern Caucasus
Lenta.ru: Россия: Общество: Кручу-верчу, ничего не хочу	10.11.2013	Discussion of the national league table of universities and its problems.

The data reported by the police has similarly been criticized for significant manipulations by one of liberal think-tanks:

> *It is impossible to collect adequate statistics on the performance of the police force. Any data collected for intelligence purposes inevitably gets used for evaluation purposes and thus becomes subject of manipulations, if not outright falsifications. (Committee of civic initiatives (Komitet grazhdanskikh iniziativ) 2012: 15)*

Data manipulation also occurs at the ballot box. 2011 Russian parliamentary elections generated a wave of academic and journalist publications aimed at discovering traces of electoral

[16] Falsifications of economic statistics in Chine make global news headings: Lenta.ru: Экономика: Госэкономика: Эксперт уменьшил ВВП Китая на триллион долларов; Lenta.ru: Экономика: Госэкономика: Ложь китайских масштабов.

fraud. Enikolopov et al. (2012) used rigorous statistical methods to estimate the magnitude of electoral fraud in Russian elections.

Administrative values

NPM assumes a culture of public service honesty as given (Hood 1991: 16)

Integrity of civil servants is among central themes of this study. A whole body of literature exists devoted to issues of public sector integrity and ethics (Hunt 1995; Hutton and Massey 2006; Keraudren 1995; Kernaghan 2003). Hood (1991) places "the culture of public sector honesty" among other values that commonly appear in debates on administrative design. Administrative values are differentiated from political values. "Administrative values… relate to conventional and relatively narrow ideas about 'good administration' rather than to broader ideas about the proper role of the state in society" (Hood, 1991:10). The same administrative system may cater for changing sets of political values. For example, "equity values could perfectly well be programmed in to the target-setting and performance indication process, if there was strong enough political pressure to do so" (Hood, 1991:10).

Hood (1991:10) identifies three "families" or "clusters" of administrative values: *sigma, theta* and *lambda*: "[b]roadly, the 'sigma' family of values relates to economy and parsimony, the 'theta' family relates to honesty and fairness, and the 'lambda' family relates to security and resilience". Emphasizing different types of values leads to differences in the focus of policy design and implementation:

Lambda-type values. "If lambda-type values are placed at centre stage, the central concern is to avoid system failure, 'down time', paralysis in the face of threat or challenge. Classic expressions of lambda-type values include: redundancy, the maintenance of back-up systems to duplicate normal capacity; diversity, the maintenance of quite separate, self-standing units; robustness, use of greater amounts of materials than would ordinarily be necessary for the job" (Hood, 1991:14).

Theta-type values. "[If] theta-type values are placed at centre stage, the central concern is to ensure honesty, prevent 'capture' of public bodies by unrepresentative groups, and avoid all arbitrary proceedings. Classic initiatives promoting theta-values include: recall systems for removing public officials from office by popular vote; 'notice and comment' and 'hard look' requirements in administrative law; independent anti-corruption investigatory bodies, freedom of information laws, extensive public reporting requirements" (Hood, 1991:13). The major concern of policies oriented at promoting theta-type values concentrate on "dishonesty and abuse of office" that lead to "palpable waste of resources", but other aspects of public sector ethics such as "less tangible stakes, notably public trust and confidence and the ability to exercise citizenship effectively" (Hood, 1991, 13). Hood remarks that "where honesty and fairness is a primary goal, the design-focus is likely to be on process-controls rather than output controls", "'Getting the job done' in terms of aggregate quantities is likely to be supplemented by concerns about *how* the job is done…concern with process may cause the emphasis to go on the achievement of maximum *transparency* in public operations" (Hood, 1991:13).

Sigma-type values. "[if] sigma-type values are emphasized, the central concern is to 'trim fat' and avoid 'slack'". Typical examples of initiative promoting sigma-type values are: "'just in time' inventory control systems, payment-by-results reward systems and administrative 'cost engineering'" (Hood, 1991:12).

Hierarchy of values

Regarding the compatibility of these three sets of administrative values, Hood (1991:15) remarks that although there may exist an overlap between them, "for example, dishonesty frequently creates waste and sometimes leads to catastrophe" (ibid), "it is hard to satisfy all three value sets equally" (ibid). There is potential for conflict between these three sets of values. Managing by results within the doctrine of New Public Management is linked to shifting away from process to output controls and greater freedom of managers. Theta-type policies, in contrast, focus on "how the job is done" and emphasize process controls and procedures.

Efficiency and economy (sigma) may come in conflict with the necessary redundancy and duplication necessary for reliability and resilience in the face of unforeseen threat (lambda).

It may be inferred from Hood's (1991) reference to theta and lambda-type values as "capital base", that the three clusters may be schematically represented by a pyramid, though Hood himself does not make such representation (Figure 8).

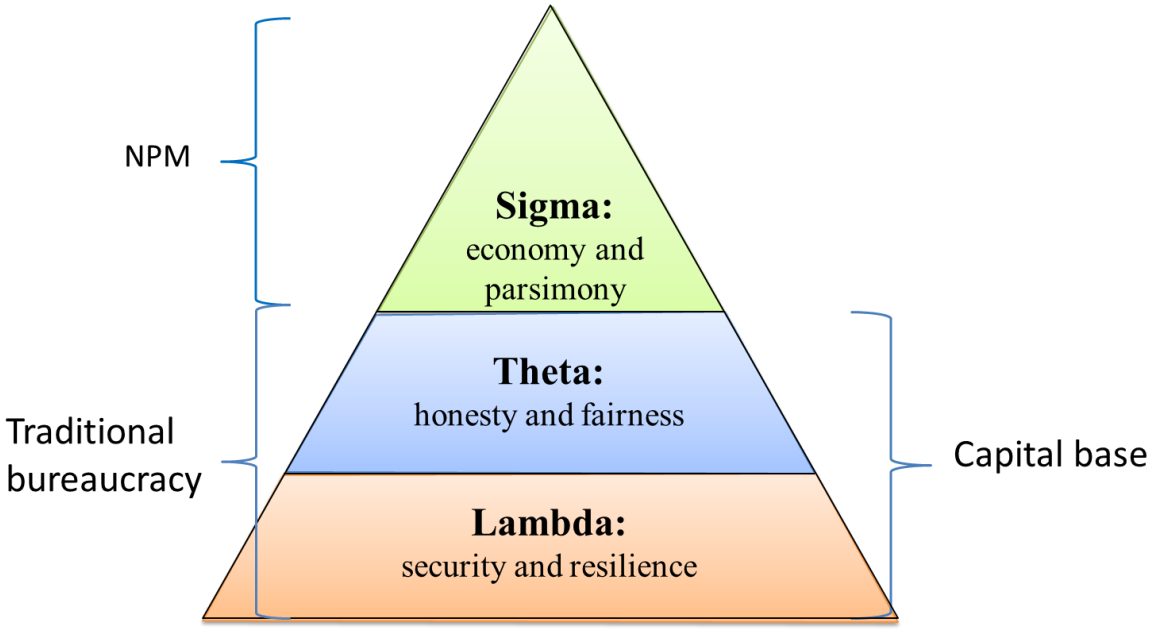


Figure 8. Hierarchy of administrative values. Based on Hood (1991)

NPM, as an administrative philosophy assumes lambda and theta values as given and concentrates on promoting sigma values of economy and efficiency: "if NPM is a design for putting frugality at centre stage, it may at the limit be less capable of ensuring honesty and resilience in public administration" (Hood 1991: 15). "NPM can be understood as primarily an expression of sigma-type values. Its claims have lain mainly in the direction of cutting cost and doing more for less as a result of better quality management and different structural design" (Hood 1991: 15). There have been concerns raised that promotion of sigma-type values may be achieved at the expense of more traditional values of bureaucratic ethics, such as honesty, transparency and democratic representation: "the extent to which NPM is likely to induce corrosion in terms of such traditional values remains to be tested" (Hood 1991: 16).

NPM originated in developed Western economies as a response to specific challenges faced by governments of these countries, including demands from growing mass of "white-collar, socially heterogeneous population" constituting the backbone of a mature civil society (Hood, 1991:7). "NPM assumes a culture of public service honesty as given" (Hood 1991: 16). Hood considered traditional theta values the "capital base" of ingrained public service culture" and highlighted the deficit of such capital base in Eastern European countries, thus anticipating the clash between "NPM clones diffused by public management 'consultocrats'" and contexts with little capital base of established public sector ethics (Hood 1991: 16).

In the literature on NPM, the emphasis is generally put on the usefulness of performance indicators for operating managers (Behn 2003) . This is in line with Hood's assertion that NPM focuses on sigma values while assuming theta values as already established. This illustrates the fact that even among the researchers of NPM in the West the danger of overlooking theta values and disregarding their development has been recently recognised. If something is neglected and falls into disrepair and disuse, it may gradually wear out and fade away. This may happen to the traditional values of integrity and other values that characterised the public sector culture in a traditional bureaucracy.

Hood made his predictions at the dawn of NPM conquest of the global public administration scene. A decade later calls for greater attention to theta values became pronounced: Barrett (2004: 261) warns against leaving managers alone to manage, as they may be susceptible to misconduct and abuse if *theta*-values of integrity, loyalty and serving the public good are eroded. One way of revitalising public sector ethics may be via ethical codes: "increased attention to ethics and social responsibility in the policy process is overdue, and would bring into the implementation debate issues such as the value conflict between professional principles and codes of ethical practice versus the management performance imperatives" (Barrett 2004: 261). The overemphasis on *sigma*-values has been seen as one of drawbacks of NPM. (Piotrowski and Rosenbloom 2002).

Conclusion

Existing literature provides evidence for a simple hypothesis: performance indicators are more effectively used if they emerge from within the organization rather then from outside. If a system of indicators is imposed by a supervising authority the risk of it degenerating into empty bureaucratic exercises is greater.

It seems that routine publication of performance data is not sufficient to guarantee citizens' control over government. Citizens appear to be interested in performance figures only when some scandal breaks out. Similarly, politicians exhibit only sporadic interest in performance figures. One may conclude that performance management largely remains an activity of managers and for managers.

Recent articles question effectiveness of performance management reforms. It may be that what we are witnessing is a natural stage in the life cycle of any managerial idea: after a big bang of interest performance management seems to have gone through its maximum popularity and is now loosing vitality.

This is the situation in developed Western economies. Compared to a large number of published articles on Western economies little is known about the implementation of performance measurement in non- or semi-democratic contexts. Russia is virtually absent from the international debate on performance measurement (excluding references to anecdotes from the Soviet period). This study attempts to bridge this gap by exploring the recent developments in implementing performance management in Russian.

CHAPTER 4. METHODOLOGY

This chapter begins by reviewing ontological and epistemological assumptions of the current study. It then progresses to outline data and methods used at different stages of this research project. It concludes by describing five major steps of the research.

Ontological, epistemological and methodological assumptions

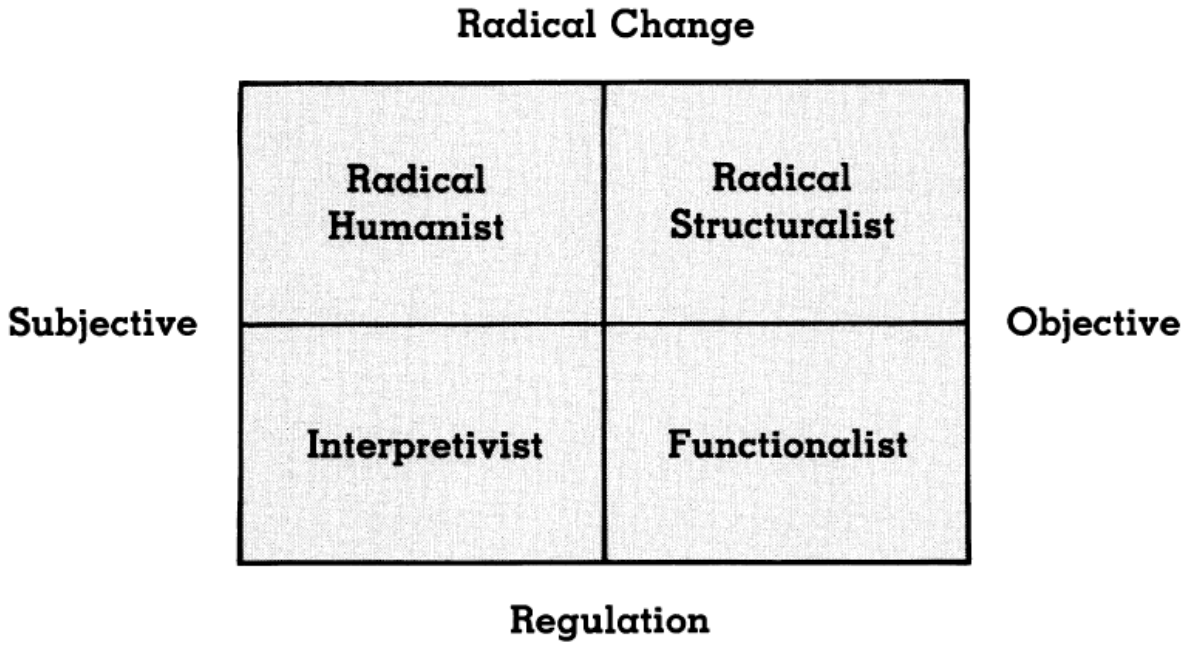
This research project takes multi-paradigm approach to study of organizational phenomena. The multitude of organizational being cannot be adequately accounted for from within any single research paradigm: «[b]y now … the field recognizes that the use of any single research paradigm produces too narrow a view to reflect the multifaceted nature of organizational reality» (Gioia and Pitre 1990: 584).

To characterize the paradigmatic approach taken in this research I use the typology of research paradigms originally developed by Burrel and Morgan (1979). This typology is based on categorizing studies of organization according to their assumptions regarding the nature of the phenomenon of organization (ontology), the nature of our knowledge of this phenomenon (epistemology) and the ways of studying this phenomenon (methodology):

> *This debate is perhaps most succinctly characterized according to differing fundamental assumptions about the nature of organizational phenomena (ontology), the nature of knowledge about those phenomena (epistemology), and the nature of ways of studying those phenomena (methodology) (Gioia and Pitre 1990: 585).*

The typology is a 2x2 matrix identifying four research paradigms: radical humanism, radical structuralism, interpretivism and structuralism. Horizontal axis represents ontological and epistemological assumptions (left – subjectivism; right – objectivism). Vertical axis represents methodological assumption and aims of the study (top – the aim is critique and radical change, bottom – regulation, gradual improvement and maintaining status quo).

Figure 9. Four research paradigms in studies of organisations. (Burrell and Morgan 1979)



According to this typology *functionalism* is defined as a paradigm characterized by the understanding of organizational reality as something objective, existing "out there" and waiting to be discovered. The aim of research within this paradigm is to achieve gradual improvement of organizational practice. *Interpretivism* is characterized by the idea that there is no one organizational reality independent of the interpretations of social actors. There are only multiple interpretations given by social actors to their actions. The aim of research within this paradigm is understanding, maintaining status quo and constructing narratives that take into account multiple points of view. *Radical humanism* takes a subjective stance to the issue of the nature of organizational reality similarly to intepretivism. However, in contrast, the main aim of research within this paradigm is to change socially constructed rules of interaction of individuals and liberate the individual. Within *radical structuralism*, the organizational reality is seen as having its objective existence (there deemed to exist structures underlying human interactions and shaping organizational practice, these structures may be discovered and changed). The aim is to change organizational practice through changing these underlying structures.

Gioia and Pitre (1990) observed that most studies in the field of organization theory tended to take the functionalist approach. They linked this to the influence of natural sciences:

> *Organizational science has been guided pre-dominantly by the assumption that the nature of organizations is a basically objective one that is "out there" awaiting impartial exploration and discovery (Gioia and Pitre 1990: 586)*

However, the objectivist position, characteristic of natural sciences, does not allow to account for many important aspects of organizational phenomena. The subjective nature of organizational reality becomes more widely accepted today and there is a growing emphasis on understanding the process of change, not just on explaining existing differences (Kvale 1996).

Gioia and Pitre (1990) advocate the use of multi-paradigm approach: each of the four paradigms may be used to provide partial understanding of organizational phenomena, but a meta-paradigm approach may allow for deeper and broader understanding. The author's position is illustrated below. The borders between paradigms are blurred symbolizing the possibility of multi-paradigm research. The oval symbolizes the point of view of a researcher, who can benefit from using all four paradigms.

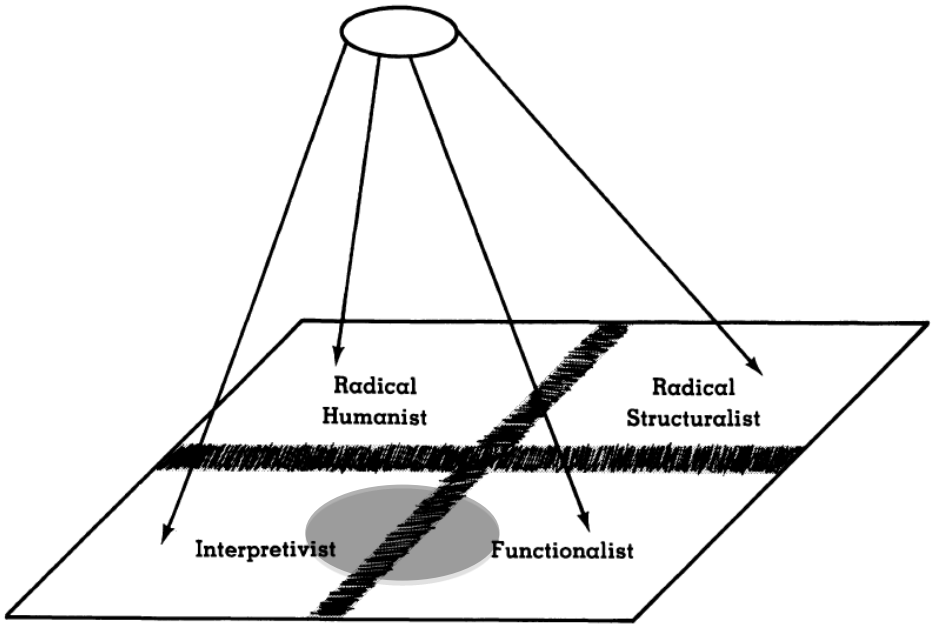


Figure 10. Meta-paradigm approach to organisation research (Gioia and Pitre 1990: 597). The grey area represents the approach taken within this project.

Willmott (1993) highlights the potential of combining research paradigms in a single project. Lewis and Grimes (1999) give a broad list of "exemplary" studies that utilized various combinations of research paradigms in researching organizations.

Mixed methods are becoming more popular in the field of organizational studies (Bryman 2006b). The "paradigm wars" of 70-s and 80-s have come to an end (Bryman 2006a) and the practice of combining qualitative and quantitative methods in a single study has becoming more and more acceptable. The pragmatic attitude that "whatever works" is acceptable is becoming widely accepted (Bryman 2006b).

The approach taken in the current study may be classified as a combination of interpretivist and functionalist paradigms (the grey area at the picture above). This combination was chosen because of its potential in theory building when a new phenomena is research or a new research method is tried (Eisenhardt, 1989). Interpretivist approach allows for new problems to be brought to the researcher's attention and for new theories to be built based on social actors' accounts. Functionalist approach is useful in testing theories and obtaining generalisable results. Thus, the combination of the two approaches allows for a full cycle of research to be carried out – from building a theory to empirically testing it.

Positioning the research along the objectivism-subjectivism continuum

The opposition of subjectivism and objectivism is not a dichotomy, but a continuum that represents different ontological and epistemological assumptions (Morgan and Smircich 1980). The two grey areas on the picture below represent two positions taken in the two stages of this project (Table 8).

Table 8. Objectivist – subjectivist continuum in social sciences (Morgan and Smircich 1980: 492).

	Subjectivist Approaches to Social Science					**Objectivist Approaches to Social Science**
	Stage 1				Stage 2	
Core Ontological Assumptions	reality as a projection of human imagination	reality as a social construction	reality as a realm of symbolic discourse	reality as a contextual field of information	reality as a concrete process	reality as a concrete structure
Assumptions About Human Nature	man as pure spirit, consciousness, being	man as a social constructor; the symbol creator	man as an actor; the symbol user	man as an information processor	man as an adaptor	man as a responder
Basic Epistemological Stance	to obtain phenomenological insight, revelation	to understand how social reality is created	to understand patterns of symbolic discourse	to map contexts	to study systems, process, change	to construct a positivist science
Some Favored Metaphors	transcendental	language game, accomplishment, text	theater, culture	cybernetic	organism	machine
Research Methods	exploration of pure subjectivity	hermeneutics	symbolic analysis	contextual analysis of Gestalten	historical analysis	lab experiments, surveys

The table developed by Morgan and Smircich (1980) was updated by Cunliffe (2011) to reflect recent developments in the philosophy of science. It is useful for the current project to note the additional continuum added by Cunliffe: the role of the researcher changes depending of the research paradigm (Table 9). If research is seen as science, then traditional criteria of quality of scientific work are applicable. However, if research is seen as craft, as is the case with more subjectivist forms of research, then a different set of criteria is applicable. The criterion of reliability is not applicable for subjectivist studies, as the researcher himself becomes the research instrument. The criterion of statistical generalizability is replaced with theoretical generalizability the more subjectivist the research is, it is dropped altogether as research becomes more subjectivist. The emphasis shifts from external validity to internal validity in more subjectivist types of research. Kvale (1996) metaphorically describes the two possible roles of a researcher as a "miner" or as a "traveller":

> *…the miner metaphor pictures a common understanding in modern social sciences of knowledge as "given." The traveller metaphor refers to a postmodern constructive understanding that involves a conversational approach to social research. The miner metaphor brings interviews into the vicinity of human engineering; the traveller metaphor into the vicinity of the humanities and art (Kvale 1996: 5).*

Table 9. The role of the researcher in different research paradigms (Cunliffe 2011: 654).

	Intersubjectivism		*Subjectivism*		*Objectivism*	
Core ontological assumptions of research methodologies *(The nature of social reality)*	Social reality relative to interactions between people in moments of time & space. Relationally embedded. Social community.	Socially constructed realities, emerging, objectified, & sometimes contested in the routines & improvisations of people. Context is human action & interpretation.	Reality as symbolic & linguistic meanings & interpretations. Contextualized in a social site.	Discursive realities constructed by discursive & non-discursive practices & systems. Contested & fragmented. Discursively contextual.	Reality as process: interrelated actions, elements, structures, and systems. Generalizable or context-dependent.	Reality as concrete structures & behavioral patterns, subject to rules & laws. Structural integration or disintegration. Naive realism.
	← *Research as Craft*			→		*Research as Science*
Assumptions about human nature *(How we relate to our world)*	Humans as intersubjective, embodied, relational, & reflexively embedded.	Humans as intentional & reflexive subjects, constructors & enactors of social realities within linguistic conventions or routines. Storytellers.	Actors, interpreters, sensemakers. Choosing linguistic resources, managing impressions.	Humans as subjectivities, products of discourse, contested & conflicted discursive sites.	Humans as an element in the process, adapting to & sometimes managing elements. Information processors & network coordinators.	Humans determined by their environment, socialized into existing social & institutional practices & requirements. Characterized by traits etc.
		Stage 1				Stage 2

The two grey areas represent the two stages of the current project.

According to a subjectivist perspective "social reality *is* the symbolic world of meanings and interpretations. It is not some 'thing' that may be interpreted in different ways; it is those interpretations" (Blaikie, 2000, 116). This approach requires distinct epistemology. As multiple social realities are equally meaningful, the way to understand these realities – is "to enter the everyday social world in order to grasp these socially constructed meanings... Knowledge of this reality is produced by 'immersion' in it" (Blaikie, 2000, 116-20).

Interpretivism and functionalism

Table 10. Aims of research and the use of theory in interpretivist and functionalist research (Gioia and Pitre 1990: 591).

Interpretivist Paradigm	Functionalist Paradigm
Goals To DESCRIBE and EXPLAIN in order to DIAGNOSE and UNDERSTAND	**Goals** To SEARCH for regularities and TEST in order to PREDICT and CONTROL
Theoretical Concerns SOCIAL CONSTRUCTION OF REALITY REIFICATION PROCESS INTERPRETATION	**Theoretical Concerns** RELATIONSHIPS CAUSATION GENERALIZATION
Theory-Building Approaches DISCOVERY through CODE ANALYSIS	**Theory-Building Approaches** REFINEMENT through CAUSAL ANALYSIS

Gioia and Pitre (1990) summarize the differences between interpretivist and functionalist paradigms in a concise table: they may be contrasted in terms of the aims of the research and approaches to theory building.

The aim within interpretivist paradigm is to *describe* and *explain* organizational phenomena in order to *diagnose* and *understand* them. The aim within functionalist paradigm is to *search* for regularities, *test* hypothesis in order to *predict* and *control* the phenomena.

In line with these aims the process of theory building is taking place. The interpretivist approach relies on the analysis of social actors' accounts (of their interpretations of organizational phenomena). The main means of this analysis is interviewing, content analysis and coding of transcripts.

In contrast, the functionalist approach is concerned with finding correlations between variables. As a result of the difference in ontological and epistemological assumptions of the two paradigms the role of the researcher is different, too.

An interpretivist researcher tries to record social actors' accounts without introducing his own bias. He adopts the position of a "dialogic facilitator", who attempts "to minimise [one's] authorial bias by letting the natives speak for themselves as much as possible. The aim is to produce a 'polyphony' of voices rather than a single voice, in order to reduce bias and distortion" (Fontana, 1994, p.214). Since *understanding* is the main objective, it follows that abductive

research strategy is used: “the objective of understanding is the exclusive preserve of the abductive strategy” (Blaikie, 2000, 124). In contrast, a functionalist researcher places himself outside of the organizational reality in a position of an “impartial observer”, who attempts to find the truth about the characteristics of the phenomenon.

The use of theories

Within the functionalist paradigm theories are used to identify relevant variables and predict relations between them. Thus, theories are used at the very beginning of a functionalist study. In contrast, the use of theories within interpretivist is mostly restricted to later stages of the research process to conceptualise findings and provide critique of social actors' accounts. The particular approach taken in this study may be regarded as a particular branch of abductivist strategy that argues that "there [is] more to social reality than the participants' understanding of it... Therefore, it [is] legitimate to critique the participants' accounts, possibly from within some theoretical perspective" (Blaikie, 2000, 258).

Stages of the research

The two research paradigms also differ in what is considered the typical sequence of steps of the research process (Table 3) (Gioia and Pitre 1990). The first stage of this study was carried out within interpretivist paradigm, the second stage may be classified as functionalist. The fist stage may be labelled as "the understanding stage", the second – as "the explanatory stage". The results of the first stage became the starting point for the second stage.

The whole research project may be broken into five steps (Figure 11). A detailed description of these steps is given in the next section. The second (qualitative) step became at the same time the final of the fist stage and the beginning of the second stage.

Table 11. Sequence of research steps typical for interpretivist and functionalist paradigms (Gioia and Pitre 1990: 593).

Interpretivist Paradigm	Functionalist Paradigm
Opening Work SELECTING A TOPIC: What are the issues? What are the research questions? DESIGNING RESEARCH: What are data? Where to find data? How to record data?	**Opening Work** SELECTING A TOPIC: What are the issues? What are the research questions? REVIEWING LITERATURE: What do we know? FINDING A GAP: What is missing? PUTTING A FRAME-WORK TOGETHER: What are the relevant theories and variables? FORMULATING HYPOTHESES DESIGNING RESEARCH: What are data? Where to find data? How to measure data?
Data Collection IDENTIFYING SPECIFIC CASES QUESTIONING INFORMANTS: according to what is relevant to them in context	**Data Collection** PROBING REPRESENTATIVE SAMPLES OF SUBJECTS: according to the hypotheses formulated
Analysis CODING: Provide a description at the first and sometimes at second level of abstraction FORMULATING CONJECTURES: Identify the relations between concepts at first level or across levels of abstraction EVALUATING CONJECTURES: Validate with informants through new data collection FORMULATING THEORY: Identify the emerging concepts and relationships REVIEWING LITERATURE: Identify what was already known	**Analysis** TESTING HYPOTHESES: Evaluate the significance of the data according to initial problems and hypotheses
Theory Building WRITING UP A SUBSTANTIVE THEORY: Show how it all fits together	**Theory Building** WRITING UP RESULTS: Show how the theory is refined, supported, or disconfirmed Show what it tells the scientific community and the practitioners

The two research processes given in the table (Table 11) may be "stitched together" at the level of "Reviewing literature". This stitched sequence will adequately describe the current study.

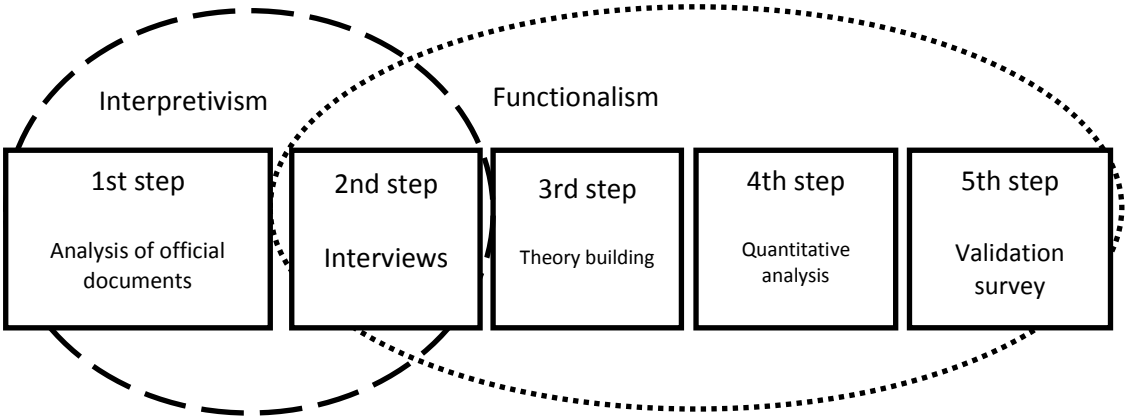


Figure 11. Steps of this research project and corresponding research paradigms.

This mixed method approach was instrumental in achieving greater internal validity and providing in-depth contextualization (by collecting qualitative data using in-depth interviews). It also allowed for practical results to be obtained and generalized both statistically and theoretically.

The second step (interviews) may be included in both stages. On the one hand, during the interpretivist stage an attempt was made to understand motives and reasons given by social actors to their actions. A discourse surrounding performance indicators was captured, this allowed 1) for in-depth contextual descriptions to be produced and 2) to understand what factors social actors regard as important in determining the way they use performance indicators.

On the other hand, the same interviews allowed for important concepts (variables) to be identified. Relations between prominent topics were examined during the interviews. Thus, the same step became part of both interpretivist and functionalist research processes.

Data and Methods

The research was conducted in two stages: first, 25 semi-structured interviews were carried out with officials from regional governments, consultants and academics. Motives and opportunities for deliberate data manipulation were identified and hypotheses regarding their

manifestation in resulting performance data were formulated. Second, quantitative evidence was obtained by analysing the nation-wide performance dataset that provides figures for over 300 performance indicators covering 83 regional governments for the period of 5 years between 2007 and 2011 (this gave over 90 000 observations).

Qualitative stage

The aim of the qualitative stage was to create a rich description of the context in which the system of performance measurement operates in the public sector in Russia. It was aimed at exploring potentially productive topics and new unexpected issues and problems existing in the Russian context. The process of interviewing began with using general questions aimed at evaluating the overall atmosphere and organizational environment in the public sector. As the process progressed, through a number of iterations topics of high relevance were identified and the scope of interviews was narrowed to focus on problems of reliability of the data and issues of data manipulation.

Selection of respondents

The sample included: 16 current regional civil servants, including three ministers of economic development, three deputy ministers from departments of health and economic development, 6 heads of departments and 2 specialists (lower level civil servants); 9 former civil servants, including: 3 consultants, 3 academics. Other respondents included: a top central government official responsible for the design of the nation-wide system of performance management and two civic activists involved in public scrutiny of government performance reports. Purpose sampling and snowballing were used to generate the list of respondents. People responsible for implementing and operating performance measurement systems were identified in regional governments and asked to provide further contacts for interviews. Snowballing began by contacting senior civil servants in regional governments who had good working relations with the researcher's institution. This approach was used because it was believed that 1) top officials

would be in a position to have a broader picture of the working of the organisation as a whole and would be able to provide relevant contacts, and 2) this was more practical in terms of the ease of getting access. Taking into account the strict hierarchy of government departments and reluctance of middle-tier managers to participate in research without formal instructions from their superiors, the only practical way of getting access was to approach it "from the top". Moreover, informal endorsement of the researcher by top-level officials created an atmosphere of trust between interviewees and the researcher. Respondents felt more comfortable to share their thoughts about issues they faced when dealing with performance measurement, if they felt that the researcher was an insider who was authorized to research the issue.

Academics were selected for interview because many of them had previous experience of working for the government either as civil servants or as consultants. Moreover, academics are considered among stakeholders in the reform discourse. They may benefit from the dissemination of a particular management idea that "provides a clear stimulus to academic production and a focus for academic debates – in short, for professional talk" (Pollitt 2001: 942) and "may offer a passport into consultancy" (ibid).

Interviewing procedure

The interview guide from (Pollitt 2006a) was used with minor alterations (the last question was removed as repetitive and Section E. was added with one question about external factors influencing performance indicators) (See Appendix 1). Questions were translated into Russian (See Appendix 2). The list of questions was used to structure the interview, but additional questions were asked to clarify and explore respondents' answers in greater detail. Interviews were recorded and transcribed. After the first round of interviews (10 interviews) transcripts were analysed and main emerging themes were identified. At this stage some questions were marked as unproductive, as it became apparent that a saturation point was reached. In subsequent interviews they were only briefly raised to check if the respondent confirmed earlier results. The decision was made to explore the issue of data manipulation in greater detail and a set of

questions were added to the interview guide dealing with the problem of data manipulation and trust in performance data. Further 10 interviews were conducted, recorded, transcribed and analysed. The remaining 5 interviews were conducted with academic researchers and consultants familiar with performance measurement practices in public sector in Russia. During these interviews findings related to data manipulation were discussed to validate them. Respondents were asked to give their interpretation of findings and provide comparisons or parallels with their own experience to validate or invalidate the findings.

Quantitative stage

Data source

The data were collected by the Ministry of Regional development in accordance with the Presidential Decree № 825 of 28.06.2007 "On assessing the effectiveness of executive authorities of the subjects of the Russian Federation." The data were published in 2012[17].

Characteristics of the dataset

The original dataset had the following characteristics:

- Number of indicators reported: 295 in 2007, 325 in 2011.
- Number of regions: 83
- Time period covered: 2007-2011 (actuals), 2012-2014 (planned)
- Observations in initial dataset: 151152 (actuals only included)

Growth indices were calculated on the basis of actual values for the period 2007-2011. The resulting dataset had the following characteristics:

- Number of indicators reported: 295 in 2007, 325 in 2011.
- Number of regions: 83
- Time period covered: 2008-2011 (grown rates)
- Number of observations: 121072

[17] URL: http://www.minregion.ru/upload/documents/2012/10/101012/101012_itogi_2011.xls [accessed 08.03.2013]

- Number of useful observations: 96107 (growth rates could not be calculated if data for one of the years were not reported).

Method of data analysis

Descriptive statistics and visualization were used to analyse the distribution of growth indices.

Official use of performance data

The data were used by the Ministry of regional development to allocate grants to regions[18]. At the same time, these data were also used by the presidential administration to evaluate governors[19] and by governors to evaluate their subordinates. Regional administrations used the same data to evaluate municipal civil servants. Thus, a chain of principal-agent relations was created within the "vertical line of power". The same data were used multiple times by multiple principals to evaluate their corresponding agents.

Validation survey

Rationale for conducting the survey

The main aim of the survey was to attempt to estimate the "natural" variation of performance indicators. Since I did not know the principle behind allocation of performance indicators to regional governments or other bodies, I could not estimate the effect of this selection on variation in the dynamics of indicators. The conclusiveness of my evidence of data manipulation significantly depends on accounting for alternative explanations of causes of the difference in dynamics between two groups of indicators. It is possible that the observed effect

[18] 20 top performing regions received a total of 2 bln roubles between them to encourage better performance. From 2008 onwards the number of regions was reduced to 10 and the total sum of grants cut by half (1 bln rubbles was distributed annually in 2008-2011). The ranking criteria were radically changed three times between 2007 and 2011 creating a situation of high uncertainty. Despite this, the list of leaders was virtually unchanging between 2009 and 2011 leaving most regions out of the grant game. Most civil servants commented that they did not consider grants a significant stimulus to improve performance consistently.

[19] Governors of Russian regional governments have been appointed by the president since 2004. There are no formal rules on what is considered good performance and what career decisions would follow if performance indicators demonstrate good or bad dynamics. The system relies on the set of informal incentives to generate competition among governors (Libman, Kozlov, and Schultz 2012).

of excessive concentration around zero was the result of some factors other than deliberate manipulation on behalf of regional governments. It is, for example, possible that indicators reported by regional governments were naturally less variable and tended to concentrate around zero without any deliberate intervention from civil servants. Perhaps, the underlying field of public administration was naturally hard to change. This could be one of the reasons why these indicators were allocated to one of the groups and not to the other. I was not able to establish the reasons behind the decisions to allocate different indicators to different groups (this would have required interviewing people in central government to which I had no access). I thus could not find out whether "natural" variance played a significant role in this selection. I could not know the natural variance of any given indicator (free from manipulations), because all I had was the data that (potentially) included manipulations.

Not being able to estimate the "natural" variance "from the top" I attempted to estimate it "from the bottom". An attempt was made to evaluate the natural variance of indicators by conducting a survey of civil servants who could give estimation (an expert judgment) of various characteristics of these indicators. Admittedly, this was not the most theoretically sound way of trying to estimate the natural variance of the indicators, but it was the only route practically available within the scope of my project.

I carried out this estimation by including five questions in the questionnaire that was used in a wider survey carried out by the Institute of public administration of the Higher School of Economics.

The survey was aimed at municipal servants. This, although, again, not ideal for my purposes, was the only practically available option. No comparable survey could be carried out on regional civil servants because of problems of getting access.

The overarching idea behind the survey was that by asking civil servants to evaluate hypothetical behaviour of performance indicators (in the absence of government interventions or in a situation of unlimited available resources) it would be possible to (however imperfectly)

arrive at some measure of "natural" variance of the phenomena that is being measured. This estimation of "natural" variance could then be compared with actual – post-manipulation – variance in performance data.

Selecting indicators for evaluation

I maximized comparability with my original estimations by including in the survey those indicators that were reported by both regional and municipal authorities. By doing this I ensured that the indicators that I was asking respondents to characterize were both familiar to them (since they were used in performance reports of municipalities) and relevant to my analysis (since they were also used in reports of regional governments to the central government).

In addition to 40 indicators that were originally used in the municipal survey, I was able to add 10 more indicators that were not currently used by municipalities, but were used in the past, or were similar to the one they were using.

I sampled indicators using quota-sampling to ensure that indicators from all three groups (self-reported, reported by federal statisticians and by federal ministries) were included. I also ensured that both indicators that demonstrated suspicious dynamics and those that demonstrated "normal" dynamics were included. The measure of "suspiciousness" was calculated as the total share of observations in the interval between -0,5% and +0,5% of annual growth. Thus, indicators that had anomaly high clustering of values around zero were marked as suspicious.

Methods of administering the survey

The survey was sent to key contacts in 5 regional governments who were responsible for cooperating with the HSE on this survey. These 5 regional civil servants distributed the link to the survey to all municipalities in their regions. Answers were received from municipalities in 3 of the regions. In these three municipalities out of 128 municipalities completed questionnaires were received from 64 municipalities (response rate 50%). Additionally, partially completed answers were received from 56 municipalities. The survey consisted of 3 separate sections that

could be completed independently. Out of the maximum of 384 sections, 258 were completed (some municipalities only completed one or two sections), giving the overall response rate of 67%. Such high response rates may be explained by the fact that municipalities were mandated to complete the survey by their superiors in regional governments. The survey was part of a wider initiative of getting feedback from municipal governments on the implementation of performance measurement procedures. This initiative was driven and controlled by the central government (Ministry of regional development) and municipalities were required to participate. The Higher school of economics provided consultancy and technical assistance to the central government in carrying out this exercise.

Structure of the questionnaire and instructions to respondents

The questionnaire was broken into three sections, each corresponding with major fields of public/municipal administration: 1) Economy; 2) Social Services; 3) Local government. The survey was administered online. Municipal servants received the link together with a cover letter that included instructions. Each of three sections could be completed independently by a person most familiar with the field. It was suggested that a person who is responsible for compiling and reporting performance data be the one who completes the questionnaire. If this person lacked knowledge to provide answers to questions in one of the sections, it was recommended that s/he should seek advice from another civil servant who is better acquainted with the field.

Interpretations of results

Q1. *In your opinion, how significant is the influence of municipal governments over a given indicator? (0 to 100%)*

It was expected that municipalities would have greater influence over indicators for which data are locally collected. The first question served as a measure of robustness of the survey and a test of respondents' engagement with the questionnaire.

Q2. *In your opinion, if the government (local and regional) does not intervene, what would be the dynamics of this indicator during the next 3-5 years?*

This question is an attempt to estimate the "natural" propensity of indicators to converge to zero. Respondents were asked to characterize the dynamics of the indicator in a hypothetical situation where no government intervention is made. This aggregate estimation was later to be compared with the actual variance of indicators to see if indicators with similar predicted *ex ante* behaviour were behaving differently *ex post* when reported by different reporting agencies. (It is possible to imagine that respondents were aware that indicators were being reported by different agencies and were, in fact, giving their answers based on their estimation of "post-manipulation" and not "natural" variance of indicators. This possibility cannot be discarded, but is likely that municipal civil servants were not interested in such elaborate mystification).

Q3. *In your opinion, if the government had abundant resources and could allocate them to the area measured by this indicator, how easy would it be to achieve growth of 5% in a year?*

This question is an attempt to measure the natural "elasticity" of indicators, that is, the ease of demonstrating significant growth by allocating greater resources to the measured area. Similarly to the previous question, it was assumed that respondents would base their answers on their estimation of "pre-manipulation" variance of the underlying phenomena.

Q4. *In your opinion, is your municipality able to collect objective and accurate data on the value of this indicator? Do you agree with the following statement – "Our municipality can collect accurate and reliable data to measure values of this indicator".*

This question is an attempt to evaluate the perception of the quality of data available to civil servants for the measurement of performance. It is expected that federally generated data will be perceived as more trustworthy. Based on the answers during interviews, I assumed that civil servants might manipulate data because no objective actual data were available to them (if it

is too costly to collect data, the solution may be to simply put in some discretionary figures). In this case indicators with lower perceived quality of input information would be more likely to exhibit suspicious behaviour.

Q5. *In your opinion, how accurate and objective is the data on values of this indicator, reported by other municipalities in your region? Do you agree with the following statement – "Other municipalities of my region provide reliable and accurate data on values of this indicator".*

This question is an attempt to evaluate general trust in data reported by other municipalities and whether it varies based on the source of data (reporting agency). The underlying expectation behind asking this question was that civil servants were aware of potential manipulations and distrusted data provided by other municipalities. It was expected that lower levels of trust would be associated with higher levels of "suspiciousness" of observed dynamics. It was also expected that lower level of trust would be associated with lower perceived quality of available data. If one municipality cannot accurately measure one of the compulsory indicators then it has reasons to suspect that other municipalities also cannot or do not measure this indicator rigorously.

Q2 and Q3 would give an approximate evaluation of the "natural variance" of the underlying phenomena assuming that civil servants base their estimations on the *ex ante* "real life" performance and not on the "post-manipulation" data. The results would only be meaningful, of course, if civil servants really engaged with the questionnaire and did not treat it as an empty formality.

Steps of the research

The research project was broken down into five steps. During each of these steps particular aims[20] were set and different research methods were used. The outlines of the steps are summarized in the table below. (Table 12).

First step – preliminary overview of context

During this stage official documents underpinning the introduction of performance indicators were analysed. Laws, degrees, ministerial reports and other documents were collected and analysed to establish recurrent themes and major justifications that had been officially given to support the introduction of performance indicators. Current Russian public administration scholarly literature was reviewed to identify relevant debates and issues. The formal logic of the system of performance measurement in the public sector was outlined and crucial elements were selected for further investigation. The practice of implementing Decree 825 was identified as a candidate for in-depth research. Links between this practice and wider performance measurement reforms were identified. The regional level was identified as most suitable object of study both on practical and theoretical grounds.

Main historical stages in the development of the current system of performance measurement were identified and relevant literature was reviewed. At this stage official documents were analysed to establish official reasons for the introduction of formal performance measurement in the public sector. Moreover, the wider socio-political context was observed and taken into account.

[20] I used Blaikie's (2000, p.75) terminology to differentiate between such aims as *exploration*, *description*, *explanation* and *understanding*: "*Explanations* identify *causes* of events or regularities, the factors or mechanisms that produced them, and *understanding* is provided by the *reasons* or accounts social actors give for their actions" (Blaikie, 2000, p.75).

The main aims of this step were exploration and description of the existing system from the point of view given in official documents, policy papers and news coverage of official policy statements.

The following questions were posed during this stage:

1. What are the key reasons officially given for the introduction of performance indicators in the system of regional government in Russia?
2. What existing administrative instruments utilize performance indicators?

These questions may be subsumed under one general question: "How is the system of performance measurement (and indicators in particular) *supposed to function* according to the officially stated position of the government?"

Results of this research step were used to inform further steps. In particular, it was essential to acquire knowledge of official context to be able to converse with civil servants in their parlance during interviews.

Second step – qualitative evidence and description of context

During this step interviews with civil servants were conducted in regional governments. The researcher also was involved in a consultancy project in one of the regional governments. This allowed for observation of the actual working of the system of performance indicators that had been evolving since its first implementation in 2007. The aim of this stage was to produce a "thick" description of practice surrounding performance indicators. In addition to interview transcripts, field notes were also kept and became part of the data. Interviewees proceeded in three rounds. During the first round 10 civil servants were interviewed. Each interview lasted from 20 to 60 minutes.

Snowball sampling was used. First, a key person (usually a minister of economic development or his deputy) were contacted and interviewed. He was then asked to provide further contacts of people in his or other departments who could be interviewed. The procedure was reiterated with the next interviewee..

Transcripts were analysed using methods of open and axial coding to generate topologies, categorize and classify recurring events, actions, problems, suggestions. Connections between these categories were sought. Then a "core category" was selected for further investigation (Blaikie, 2000, 239). On the other hand, questions for which the saturation point was reached were identified (Eisenhardt 1989). During this stage the problem of data manipulation was selected as the "core category" for further research and the scope of the study was narrowed to only include performance indicators collected according to Decree 825. All other performance measurement procedures were regarded as context and not the object of the study. During the second round of the interviews (10 more civil servants) original broad questions were only briefly raised to elicit respondents' overall position. In this way more time was left to deeper exploration of questions related to the core category of data manipulation. In the third round of the interviews, after another round of analysis, research findings were discussed with government consultants and academics whose research interests included performance measurement. Additional insights, examples and judgements regarding validity of the findings were sought.

In addition to conducting interviews the researcher also attended public hearings on the issues of strategic planning in one of the regions and participated in consultancy projects that were carried out by the Higher School of Economics for regional governments. Observations and notes also were included in the data for analysis. At the public hearings two local activists were interviewed to obtain their views on the effectiveness of performance measurement system of regional governments. Transcripts of speeches made at public hearings were included in the dataset.

The overall question raised during this step may be formulated in the following way: "How does the system of performance measurement (and performance indicators in particular) *actually function* at the regional level of Russian government?"

More specifically, the following research questions were posed:

1. From civil servants' point of view, do performance indicators fulfil their officially declared purposes?
2. What are the actual functions of performance indicators, as perceived by civil servants?
3. What are the rewards/sanctions for over/underperforming against the indicators?
4. What are the problems with using performance indicators?

Once the problem of data manipulation was identified as the core category, more specific questions related to this issue were raised:

1. What strategies of data manipulation are used?
2. What stimulates civil servants to manipulate data?

Table 12. Research steps summary.

Characteristic	1 step	2 step[21]	3 step	4 step	5 step
Name	Preliminary	Qualitative (Interviews)	Theory building	Quantitative analysis	Validation
Stage	Understanding		Explanation		
Main type of data	Official documents (laws, governmental and presidential decrees, ministerial regulations, etc.)	Interview transcripts, field notes	Results of 1st and 2nd steps.	Performance indicators database (Presidential decree 825)	Online-survey of civil servants
Method of data collection	Internet search	Semi-structured interviews, field observations, recording and transcription	-	Secondary data from existing database	Online survey
Method of data analysis	Content analysis, overview	Coding, content-analysis	Induction, deduction	Statistical analysis	Regression analysis
Aim	*Exploration, description* of the field from legal/official point of view	*Exploration* of the existing discourse surrounding performance reporting. Identification of subject-positions and groups involved. *Understanding* of motives of actors and their interpretations of the discourse. Selection of explanatory concepts for theory building.	Models developed and hypothesis formulated. *Theoretical generalization* using Donwns' theory of bureaucratic personality types.	Hypothesis testing. Explanation of observed regularities using the newly developed theory. *Statistical generalization.*	*Validity of results tested Triangulation* Principal *falsifiability* demonstrated
Result	Legal environment regulating the system of performance	Main subject-positions in the discourse surrounding performance reporting were	The problem of data manipulation selected as the core problem for further	Hypotheses tested. Quantitative findings provided further evidence of	Results *validated* through triangulation.

[21] The zig-zaging line symbolizes overlapping between these two steps.

	measurement is *described*. Main policy documents regulating performance measurement in the public sector were identified.	identified and *described*. Based on respondents' accounts main problems related to performance indicators were identified. Motives and interpretations attributed by the respondents to other subjects and to themselves were *understood*.	research and *explanation*. Key explanatory concepts were selected for theory building: principal-agent relations in the chain of data reporting. Models of data manipulations were formulated and hypotheses regarding data patterns were derived. Theory of bureaucracy by Downs and Tullock used for theoretical generalization.	data manipulation and allowed for statistical generalization.	Potential *falsifiability* of the theory demonstrated.

Thus the three rounds of interviews allowed for the following aims to be achieved: exploration (respondents' reaction to the topics raised by the researcher was explored and new topics were identified using respondents' answers); description (attitudes of respondents were described and their interpretations of social reality were noted); understanding (main subject-positions occupied by respondents in the discourse surrounding performance reporting were identified, motives and interpretations that respondents attributed to other social actors and to themselves were understood).

Third step – theory building and theoretical generalization

The main aim of this step was to build a theory that would predict quantitative effects of data manipulation on the national scale. Using examples and interpretations provided by respondents, two strategies of data manipulation were constructed: the "prudent" and the "reckless" strategies of manipulation. On the basis of these strategies two models of data manipulation were formulated and hypotheses regarding expected quantitative effects were derived. These hypotheses were tested during the next research step on the dataset generated by the national performance measurement system.

The theory was generalized by linking it to broader theories of bureaucracy: Antony Downs's theory of bureaucratic personality types (Downs 1967) and Gordon Tullock's theory of negative selection against honesty in formal systems of performance measurement (Tullock 2005). In terms of the theory of bureaucratic personality the two strategies of data manipulation would be explained by linking them to two "self-interested" personality types of bureaucrats: the "prudent" would correspond to the ideal type of a "conserver", whereas the "reckless" strategy would correspond to the ideal type of a "climber" Theoretical explanation in terms of inherent incentives of bureaucratic hierarchies allows one to generalize findings of this project to a wide range of organizational contexts. Whenever formal performance measurement is introduced in a hierarchical bureaucracy, incentives for data manipulation appear.

Forth step – quantitative evidence and statistical generalization

Two hypotheses formulated during the previous step were quantitatively tested using performance data from the national dataset. Quantitative results provided further evidence of potential data manipulations. Even though on their own these quantitative results could not prove that performance figures were manipulated, they provided additional evidence that could be combined with qualitative findings. By combining two types of evidence (widespread scepticism of respondents and numerous cases of manipulation that they claimed to have witnessed and quantitative evidence of suspicious dynamics of indicators) it became possible to argue more persuasively that the two strategies of data manipulation really do take place and have a significant effect on overall performance statistics.

The aim of this step was to 1) describe the dynamics of performance indicators and identify indicators demonstrating suspicious behaviour, and 2) to explain suspicious dynamics of these performance indicators using the theory developed during the third step of the research.

Statistical generalization was performed by using descriptive statistics and visualization to demonstrate patterns in national performance data that correspond to predictions regarding two strategies of data manipulation. It was possible to show that cases of data manipulation are not limited to those listed by respondents but have a universal character.

This step allowed me to *verify* the theory developed at the previous step. The research question during this (forth) step was formulated in the following way: "Are predicted effects of the two strategies of data manipulation observable in the combined national performance dataset"?

Fifth step - validation through triangulation

This step was taken to demonstrate *validity* of the findings and their potential *falsifiability*. An attempt was made to validate quantitative findings by triangulating them with subjective valuations of civil servants. To achieve this, a survey was carried out. 173 municipalities were surveyed online to obtain their assessment of reliability of performance figures. 50 indicators

were sampled to include both those with “normal” and “suspicious” dynamics. During this step it was demonstrated that quantitative findings are potentially falsifiable.

The following research question was asked during this step: “Are indicators that are suspected of being manipulated, also perceived as less trustworthy by civil servants?”.

In addition, a attempt was made to obtain an independent measure of “natural variance” of phenomena measured by the sampled indicators. This was done to try to counter the main deficiency of the quantitative method of identifying effects of data manipulation. This deficiency consists in the fact that difference in dynamics between indicators may be explained not by the effects of manipulation, but by natural differences between phenomena that are measured. It is not possible to obtain any other independent evaluation of “natural variance” of underlying phenomena other than from civil servants who may provide expert judgment on the matter. If there is no systematic difference in “natural” dynamics between normal and suspicious indicators, then this would be an additional piece of evidence in favour of data manipulation as the cause of abnormal dynamics of some indicators.

Civil servants were asked to rate a sample of performance indicators in terms of reliability of the data and of the “natural” variance of measured spheres of public administration (see Appendix 2 for details of the questionnaire and see section on methods of data collection for further discussion). Indicators with suspicious dynamics were expected to be perceived as less reliable by civil servants. Participants in the validation survey were not related to participants of the qualitative stage.

CHAPTER 5. PERFORMANCE INDICATORS IN REGIONAL GOVERNMENTS: QUALITATIVE FINDINGS

This chapter aims at generating a "thick description" of the research context by reporting qualitative findings of interviews and observations. Description of this sort is necessary to create a holistic account of the practice of performance measurement in Russian regional governments. During the fieldwork stage a lot of qualitative data were generated that were not necessarily directly related to the problem of data manipulation, yet contributed to the selection of data manipulation as the main topic for investigation. The conditions of organizational environment that were observed during fieldwork directed the choice of the topic of subsequent deeper investigation and opened some options for research while showing irrelevance of other research topics. In this case, the context in which subsequent quantitative and theoretical findings were obtained itself serves as material for triangulation and validation of the results. The same quantitative findings obtained in different contexts are likely to be interpreted differently. The context largely determines which interpretations count as plausible and which should be discarded. For this reason in this chapter I attempt to convey not only particular individual respondents' accounts but also the overall impression of their attitude towards the subject of discussion and the spirit of their remarks made in informal environment preceding or following formal interview sessions. Issues that appeared most consistently are presented first, while issues on which there was no apparent consensus are presented later.

The interviews focused on problems that civil servants were facing in their practice of implementing performance measurement procedures.

Quotes in original Russian are indicated by endnotes and are given in Appendix 8.

Agencies' (in)ability to influence the indicators

All interviewed government officials, without a single exception, expressed dissatisfaction with the fact that their organizations had very limited influence over the dynamics of

performance indicators. Among the reasons that were indicated, the following three were most prominent: 1) *long-termism* - most actions taken by the government only affect the indicators in the long run, intermediate short term results are not captured by the indicators, yet rewards are tied to annual improvements, 2) *external factors* often overrun regional efforts and 3) *inconsistencies and delays in calculations of the data* by external bodies often render performance data vexatiously volatile.

Long-termism

If government actions have any effect on the indicators, this may only be observed with a considerable lag, yet rewards are tied to annual improvements*: "What the authorities are doing today will only be reflected by the indicators in 5 years. However, these indicators are used to judge our performance today" [CS-1]*. The minister of economic development in one of the regions expressed his deep conviction in the futility of using dynamics of performance indicators as a measure of organizational or individual efforts:

> *…these indicators are only useful as a snapshot [allowing initial comparison between regions]. They are only needed once to get the overall picture. Using them to measure dynamics and efforts is pointless. There are too many external factors and the effect of individual efforts cannot be traced to results. [M-3]*

External factors and expenditure commitments

Most respondents named external factors as one of the most severe impediments to the functioning of performance indicators as a measure of individual and organizational effort. External factors named by respondents varied widely: from the current global economic uncertainty, to the Soviet socio-economic legacy. In two out of three regions respondents considered their input in measurable results very modest. A deputy minister remarked: *"regional input in the dynamics of indicators is very small, and our agency's direct influence is almost negligible" [CS-1]*.

Only in Moscow area were there an air of relative confidence and a certain feeling of initiative and self-assurance. At the time of the interviews the department of economic

development of Moscow area was vigorously implementing programme-based budgeting and there was a certain enthusiasm about promises of increased effectiveness. This initiative was launched by Sergey Shoigu, the former head of the Ministry of Emergency Situations, who became governor only a month before the fieldwork was carried out. Maybe the fact that the region just got a new charismatic leader with an excellent managerial record contributed to the overall invigoration of the department. It could also be that the sheer difference in the scale of available resources between Moscow area and the two other provincial regions contributed to the feeling of confidence of public employees. At any rate, the difference with the other two regions was notable, and the impact of the regional government was believed to be stronger.

Unfortunately, it is not clear whether this feeling of enthusiasm lasted, because right after the fieldwork was completed, Sergey Shoigu left office as rapidly and unexpectedly as he took it. After having served only two months as governor of Moscow area he was appointed Minister of Defence by President Vladimir Putin. This event demonstrates how volatile and unforeseeable the administrative apparatus was during the observed period. Since governors were appointed and re-appointed by presidential discretion this turn-over of leadership was itself an important external factor often unsettling plans and forecasts of government units.

Federal decisions were listed by respondents among the most important external factors. For example, in discussing the impact of regional authorities on the dynamics of poverty indicators, a deputy minister of economic development gave the following description:

> *...poverty .... Yes, by adjusting social security measures we can have some influence over it, but the most powerful impact ... is made by federal decisions. Take pensions, for example. Pensions were significantly increased by the federal government. Plus, they have introduced subsidies towards the minimal cost of living. As a result, poverty rate has declined.* ***Our influence here was practically zero.*** *This social legislation was federal. Yes, the money is ours, but ... the system was invented and implemented by the Federation [CS-1].*

Poverty rate is, in theory, an outcome indicator. In the case when an agency is only charged with implementation, a process indicator may be more appropriate. In the Russian case,

however, respondents never explicitly distinguished between different types of indicators and never referred to one type as more appropriate than others. This suggests that the system did not make this distinction obvious or meaningful for them. The ranking procedure employed by the Ministry of Regional Development also did not break indicators down into types explicitly.

Among other important external influences respondents included large private companies that determined the economic landscape of their region:

> *Consider such an important economic indicator as "gross regional product". Local authorities, even the most powerful governor, are unable to influence it. Take the industry, for example. The owners of large businesses are not incorporated in the region; often - not even in Moscow [Deputy Minister of economic development-1].*

Here, again the respondent used gross regional product (an outcome indicator) as an example. According to respondents, federal funding decisions constituted one of the major limits of the regional government's autonomy. A large portion of regional budget was spent on funding federal mandates. Referring to this situation, respondents expressed scepticism in the possibility of a true system of performance-based budgeting.

Since a large part of expenditure is mandated by federal (and regional) laws, regional civil servants in many cases have little managerial discretion: *"we do not fund things that we are willing to fund or that are successful, we fund things that the law says we must fund*" [Deputy minister of economic development-1]. So, if performance indicators demonstrate negative dynamics, or government programmes fail to achieve anticipated results, the funding often increases instead of decreasing, since the government is required to maintain provision on a certain legally mandated level. Such situation, is, of course, not uniquely Russian, but it seems that in the Russian case there was no explicit and formalized attempt on the part of the designers of performance measurement reform to account for this and to distinguish between indicators that cover areas of regional discretion and those that measure areas of mixed responsibility of federal and regional governments.

Methods of (re)calculation

Another factor that civil servants perceived as a limit of managerial usefulness of performance indicators was the unpredictability of calculation methods. These methods were developed by the Ministry of Regional Development and were subject to frequent and drastic modifications. For example, in the period between 2007 and 2012 the methodology of ranking regions on their performance for the purposes of allocating grants was radically rewritten three times. These were not minor correction or amendments but conceptual changes in the ideology of the calculating method leading to a complete reshuffling of performance league tables. It is worth examining these modifications in greater detail.

Redesign of ranking rules

Between 2007 and 2012 the methodology of ranking used by the Ministry of Regional Development to allocate grants was changed three times. A summary of the changes was presented in a public speech by then the Minister of regional development of Russian Federation (V.F.Basargin) in 2011 at a national conference on public administration[22].

During the first two years (2007-8) the ranking was based on absolute values of indicators. This meant that the top 20 places were taken by the richer industrial and oil-extracting regions leaving no hope for poorer regions to compete for grants:

> *… during the first two years, our top 20 regions showed almost no dynamics. The leaders, naturally, turned out to be our base regions – industrial centres and oil-and-gas regions. They had serious economic base and, naturally, were socially more affluent*[23]*.* [i]

As a result, the Ministry changed the ranking methodology and decided to reward regions for *positive dynamics* of indicators. Naturally, this radically changed the ranking. Laggards suddenly became leaders:

[22] Basargin V.F., Gaidar Forum, plenary session, 2011.
[23] Ibid.

As a result, after two years of running this system, we moved to a different criterion of effectiveness and started ranking regions on dynamics of indicators. I have to say that during this period, in 2009, we were most actively criticized. This was because we turned the system upside-down (180 degrees) and the regions with the lowest base started coming first in the ranking. They had the lowest absolute values of indicators but in terms of growth rates they were the leaders. So then we got among the leaders our Northern Caucasus regions and some regions from Central Russia that we normally consider laggards (не совсем базовых регионов Центральной России).[24] [ii]

Apparently, the Ministry succumbed to the criticism and decided that this was not the kind of ranking they wanted. Thus, the system was overturned the third time and new criteria were established to try to strike the balance between the two predecessors. This time the ranking was based on a "complex evaluation", taking into account both absolute values and dynamics with certain weights:

This is why in 2010 we made some further steps to improve the system. In 2010 the ranking was based on the complex valuation that offset the disadvantages of both previous systems. And I think that in the current year (2011) we are going to continue working on improving the methodology[25]. [iii]

It can, thus, be seen, that over the period of five years the system was rather dramatically changed three times. The changes in the ranking methodology were rather fundamental. This inevitably had to undermine the consistency of evaluations and make league tables incomparable between years.

It is understandable that respondents made sceptical remarks about the usefulness of the ranking:

Interviewer: You mentioned that governors are evaluated and ranked on the basis of these performance indicators…

Respondent: These are just games that really only distract us. (Former CS, Consultant -1) [iv]

Interviewer: Was the possibility of getting a grant considered an important motivation by you or your superiors?

Respondent: No. We didn't pay much attention to these grants. It was impossible to predict your place in the ranking, it always used to come as a surprise. (Former CS, Consultant – 2).

[24] Ibid.
[25] Ibid.

Despite the fact that this speech was given at a national conference in 2011, and the interviews were conducted nearly a year later (in summer 2012), many respondents still talked about the dynamics of indicators as the ranking criteria. Perhaps, the new methodology was not clearly communicated to the regions, or, perhaps, they still considered the dynamics of indicators as the most important criterion (or, perhaps, they simply used the same name "dynamics" for the new complex criterion).

In addition to frequent changes of the methodology, there was another problem that respondents indicated. Values of performance indicators retrospectively changed because of federal recalculations of economic indexes. Federal statistics is released with a significant lag and often introduces corrections into last year's figures that radically change performance data. One respondent remarked on this with a sense of hopelessness:

> *... these numbers (performance indicators) cannot be tracked by the statistics. They are far too numerous. For example, just today we received corrections from Rosstat (the federal bureau of statistics). They estimated the size of informal economy in 2010 and increased our figures for small business by 4 billion rubbles. It had the immediate effect of worsening the dynamics of our key performance indicator: we used to have the growth of 132% between 2010 and 2011, and now it is just 20% ... growth figures for this period have plummeted. But, of course, figures for the previous period have greatly improved. Unfortunately, no one cares about previous periods.* ***Last year's figures have been completely forgotten by now****. [CS-2]*

These technical uncertainties may create a certain "organizational amnesia". It is no wonder that the minister of economic development in one of the regions in an informal conversation likened the practice of public sector management to the plot of "*The Groundhog day*", implying that there seems to be little progress from one year to the next:

> *You know, "The Groundhog day" – this is our life here. It all comes to this: we write a report about the spring-summer period, then we start preparations for the winter season and then we write a report explaining why these preparations failed. And it happens again and again every year. It only gets worse. And then there comes the sowing period again. And always there is either rain, or draught, or a flood, whatever it is, the result is always the same – the harvest is a failure.*

This conversation followed a formal meeting where a proposal was discussed to develop a strategy for the region to plan its socio-economic development for the coming 20 years. The minister lamented the lack of strategic direction in the system.

Taking into account such perceived degree of unpredictability of one's performance (and ranking progress) in the following year, it is not surprising that respondents did not consider grants of the Ministry of Regional Development as an important incentive. It is reasonable to assume that the likelihood of getting a grant in such circumstances bears little correlation with efforts of a regional government.

Lack of interagency cooperation

Another important hindrance to managerial usefulness of performance indicators lied in the lack of interagency cooperation. Respondents spoke of the current process of allocating indicators to ministries, departments and offices as excessively top-down and formalistic. The process of assigning responsibility for an indicator to a particular department was described as "jostling" (rastalkivanie), or "hanging" (naveshivanie) onto departments. This jostling was performed by the 'apparat' (the Governor's Office), while ministry officials tried to evade it by shifting responsibility on one another.

The fieldwork covered a number of public discussions on the newly developed strategy of economic development in one of the regions. The discussions were attended by representatives of a number of ministries. The authors of the strategy (external consultants and academics) proposed to introduce a few additional performance indicators related to the quality of the environment (such as air pollution, water contamination and forests land area). This proposal was vigorously opposed by some representatives of the Ministry of forestry, who argued that these indicators were not methodologically sound and would be a bad reflection of the quality of the environment. Despite the fact, that the objection was framed around methodological issues, the presenter (an external consultant) dismissed the arguments made by the civil servants and

replied that the discussion was hollow, because what the civil servants really wanted was to evade the prospect of being assigned responsibility for this indicator:

> *…We need not argue [about methodological issues], because I think that we all understand that the real problem for you is that you don't want your ministry lumbered with this indicator (чтобы вам не* навесили *это в министерство)*[v]

Apparently, this was a fair description of the situation, because no further objections were made by the civil servants. Presumably, they were satisfied that their real objection was properly understood and that, at the same time, they did not have to state it explicitly. The consultant saved them the efforts of inventing pretences (elaborate methodological issues) and at the same time assured them that their concern was understood and would be considered for what it really meant – the sign of reluctance to take on another burden of paperwork and accountability (Goffman 2006).

Respondents gave the following description of how indicators were distributed to departments and directorates:

> *… individual indicators are the responsibility of departments, and within departments they are assigned to directorates. For example our directorate [of capital investments] has as its performance indicators the amount of investments and the amount of public (budgetary) investments. (Ministry of economic development, Head of department)* [vi]
>
> …
>
> *In the DROND there is a special section on social services. I think it's the Aim number 4. Within aims there are tasks. I reckon ours is the Task 5. Our directorate has as its performance measure the ratio of disabled people to whom rehabilitation services were administered in total population. This is one of our indicators. We have others related to other tasks*[vii] *(Ministry of social security, Head of department).*

This process of formal assignment of indicators meant that they were sharply divided into "ours" and "theirs". The resulting fragmentation of responsibility had serious detrimental effects. Some issues that could have a broad effect on the region's economic situation were sometimes not resolved for many years because they required cooperation between a number of agencies. In other words, "corporate" indicators were assigned to functional units. In all three regions the

standard practice was to assign all indicators to functional units (normally, to "*upravlenie*" (directorate) – a sub-unit of a department within the ministry-department-directorate structure). No formal provisions were made by the federal legislation to encourage formal partnerships between government departments. The head of the directorate of capital investments gave an example of the length of time necessary to obtain a construction permit as an indicator that effected investment climate in the region, yet was assigned to a different ministry:

> *Who does this indicator belong to? I mean, that on the one hand, it affects our investment climate, which is the domain of our ministry, but, on the other hand, we do not own this indicator, it is not ours*[viii].

This indicator was assigned to the Ministry of construction. This rigid separation created significant problems. The Ministry of economic development had identified that the period of time it was taking businesses to obtain a construction permit was among the longest in the country. This, they thought, had serious negative consequences for economic growth, as investors preferred to invest in other regions with less red tape. The indicator of the length of time for construction permissions was, however, assigned to the department of construction of a different ministry. This department had no interest in such abstract measures as economic growth; its sole concern was to ensure compliance with construction regulations and the law. They saw no particular benefit in reducing the time of issuing permits since this would require significant effort on their side, but would result in improving indicators of a different ministry. Thus, the problem remained unsolved for more than 5 years.

When asked about potential ways of resolving the problem, respondents first mentioned the governor as the one who could coordinate inter-agency efforts, but then admitted that he often failed at performing this function. The centralized coordination mechanism (through the governor) did not always work. Problems that are of importance for a ministry were not always the same as those of importance for the governor. Time and priorities of the governor may not let him pay sufficient attention to such problems. Here is the description given by the head of the

directorate of capital investments (the Ministry of economic development) related to the failure to organize interagency cooperation on construction permissions waiting time:

> *After one of the meetings between the governor and the minister of economic development, the governor demanded to produce an action plan on lowering this indicator (waiting times for construction permits). The minister simply shrugged and remarked that this issue was just beyond his authority. (CS-1)*

According to the respondent, the governor could have intervened to resolve this deadlock, but he did not:

> *I think that the governor had to order the Ministry of construction to thoroughly investigate this problem and come up with a solution for it. He could initiate a special enquiry with members from different agencies. But he did none of these.*[ix] *(CS-1)*

Perception of non-cooperation

Respondents perceived that this situation was not satisfactory and felt that interagency cooperation was, indeed, lacking. A formal partnership was suggested as a solution, and it was lamented that no such partnership was provided for:

> *Respondent: For example, we have the disabled persons' rehabilitation rate as our indicator. Its growth depends not only on us (social services), it also depends on whether the person was employed, whether the person was able to get quality medical service to allow him to get back to work. Of course, it also depends on us, that is, on the quality of rehabilitation social services… I think, that, perhaps, it would be good if such indicators were formally assigned to multiple ministries, so that they could share the responsibility between them.*
> *Interviewer: Do you mean that now there is no such sharing of responsibility?*
> *Respondent:* ***Now everyone is responsible only for himself****.*[x]

The initiative for creating conditions for such partnerships was expected to come from superiors: from the president or the governor. But the hope for such initiative was weak:

> *Perhaps, it could be the initiative of the president to create appropriate framework for partnerships on provision of such services that depend on many agencies. Maybe even the governor could be included in this partnership to share responsibility.* ***But who would care to do this? (Но кто возьётся?)***[xi]

Horizontal cooperation

Supplementary questions were asked to explore respondents' experience with horizontal cooperation among government agencies. Respondents did not see horizontal cooperation as a feasible alternative to the centralized coordination of agencies. They tended to consider their own efforts in initiating partnerships from below as limited:

> *Well, surely we cannot be responsible for their [disabled persons'] employment. Even though we communicate and work together with [the Ministry of Labour]. But this so called interagency cooperation does not really work (Ministry of social security, Head of Department).*[xii]

The literature on horizontal cooperation indicates that "grass root" partnerships may be expected to be more sustainable than centralized mechanisms of coordination imposed from the top (Ahn et al. 2003; Dietz et al. 2003; McGinnis and Ostrom 2012; Ostrom 2004, 2005, 2010). During the fieldwork, evidence was obtained that was in line with such theoretical predictions. For example, one respondent described the following case: a newly appointed governor requested the Ministry of economic development to provide an update on performance indicators. The respondent (lower-tier civil servant with over 5 years of experience in municipal and regional government) started contacting municipalities to collect updated data on performance indicators. To his surprise, he discovered that many municipalities had not updated their performance figures for two previous years. It turned out that over two previous years former governor was not putting pressure on municipalities to engage in preparing performance reports (DRONDs), because "the interest had faded away". Thus, municipalities stopped doing this exercise, giving the following explanation: "*Well, what do you expect? They stopped demanding DRONDs from us, so we stopped updating the data*" [CS-3]. This may be taken as evidence that DRONDs and performance indicators had little internal value for these municipalities and were mainly prepared to satisfy external requirements of the governor' office. It seems unlikely that such a system could have a stimulating effect on inter-agency and inter-government horizontal cooperation.

Legal and administrative frameworks also did not seem to encourage spontaneous horizontal cooperation. A respondent gave the following explanation for the lack of initiative on their part to engage in shared efforts: "*we are not administering the budget of this programme, so we cannot take responsibility for this indicator"* [CS-2].

Overall, it appeared that respondents expected the coordination of inter-agency efforts to be performed by some central authority (the governor), but were also aware that these expectation were not being met in the current system.

Managerial literature suggests that performance indicators may facilitate interagency cooperation because they may enable different agencies to speak the same language and have a shared purpose (Kaplan 2001; Rangan 2004). It has also been observed, however, that the same set of indicators may not be appropriate for all managerial goals (Behn 2003). It appears, that the system was not contributing to promoting interagency cooperation (either because it failed to do so or was never intended to perform this function). The following dialogue with the minister of economic development of one of the regions provides clear evidence of this:

> *Interviewer: Are indicators used in any way to facilitate cooperation between different agencies?*
> *Respondent: No, they are not. Numerous issues arise. Responsibility has to be shared, credit has to be shared, budgets have to be shared and so on. No, they are not used and the issue of cooperation is very troublesome for us [CS-1].*

The potential of performance indicators as a tool of inter-agency cooperation

Using Elinor Ostrom's work on self-regulating cooperation mechanisms it is possible to suggest a number of recommendations to facilitate horizontal cooperation via the use of performance indicators. These should be aimed at creating favourable conditions for horizontal cooperation, creation of networks and encouraging initiative "from below." In the table below an attempt is made to apply Ostrom's rules of successful cooperation to interagency interaction, explanations of the principles are taken from Kuzminov and Yudkevich (2010).

Design principle	Original meaning (Ostrom, 2000)	Potential meaning in relation to performance indicators as a tool of inter-agency cooperation	Is applied now?
The presence of clear boundary rules	This principle enables participants to know who is in and who is out of a defined set of relationships and thus with whom to cooperate.	Formalized shared responsibility for an indicator. Formalized allocation of responsibility to a group of authorities. Interagency "task forces" and collaboration teams.	No
Allocate benefits proportional to required inputs	The second design principle is that the local rules-in-use restrict the amount, timing, and technology of harvesting the resource; allocate benefits proportional to required inputs; and are crafted to take local conditions into account. …How to relate user inputs to the benefits they obtain is a crucial element of establishing a fair system.	Assigning budgets to interagency task forces. Cooperative funding in proportion to inputs. Assigning joint responsibility to indicators.	No
Most of the individuals affected by a resource regime can participate in making and modifying their rules.	Resource regimes that use this principle are both able to tailor better rules to local circumstances and to devise rules that are considered fair by participants.	Involving staff in the discussion of performance targets. Using a bottom-up approach to selecting indicators.	No
Select monitors, who are accountable to the users or are users themselves	Most long-surviving resource regimes select their own monitors, who are accountable to the users or are users themselves and who keep an eye on resource conditions as well as on user behaviour.	Interagency working groups with temporal chairing responsibilities. Invitation of elected MPs, as Chairmen. Involvement of citizen committees to oversee the progress.	No
Graduated sanctions that depend on the seriousness and context of the offense	In a regime that uses graduated punishments, however, a person who purposely or by error breaks a rule is notified that others notice the infraction (thereby increasing the individual's confidence that others would also be caught). Repeated infractions attract more and more serious sanctions. The capability to escalate sanctions enables such a regime to warn members that if they do not conform they will have to pay ever-higher sanctions and may eventually be forced to leave the community.	Progress evaluation reviews. Assessment of contributions of agencies involved. Sanctions ranging from media announcements to career decision and budget cuts.	No
Cost-effective local conflict resolution mechanisms	By devising simple, local mechanisms to get conflicts aired immediately and resolutions that are generally known in the community, the number of conflicts that reduce trust can be reduced.	A formal mechanism for regular interaction of "working groups" and progress reviews. Dispute resolution within working groups. Involvement of the governor for critical issues only.	No
Minimum recognition of the right to self-organization by the authorities	This allows local users to create rules that match their local conditions, reduces risk of interference from superior authorities and legitimizes local procedures.	Freedom to manage. At least minimal flexibility of budgets at sub-organizational level and lower barriers for horizontal cooperation.	No

Perceived lack of feedback and top-down implementation of performance indicators

Another important source of problems for a managerial reform is the process of its implementation. Whether a reform is to succeed or to fail is determined by a combination of its design features and the implementation process (Barrett 2004). The success of implementation process is, in turn, largely dependent on staff's resistance to reform. The literature on public sector reform suggests that better communication, information, clarification, and involvement may often help to overcome such resistance (Barber 2007). Involvement in the process of developing a system of indicators and establishment of goals and priorities allows the "ownership" of the reform to be brought closer to employees. For example, the importance of "ownership" is stressed by guides on reform implementation in the UK (Audit Commission 1999). If employees are involved in the reform process, it is argued, they are more likely to understand and accept it. Involvement is one of the means to overcome resistance to reform (Barber 2007: 166). Likierman (1993) focuses on the role of employee engagement. According to him, "ownership" determines how seriously and responsibly indicators are used. If people do not understand how and why certain indicators are set, they tend to treat them sceptically and without enthusiasm. "Ownership" in this sense is different from accountability or responsibility for a particular indicator. In Russian practice it is assumed that the responsibility for each indicator must be assigned to a particular department or division within the public body. Someone must be held accountable for poor results. Thus, people are assigned responsibility for indicators that they do not always "own".

In the observed practice all the emphasis seemed to be on formally assigning responsibility for an indicator to some organizational unit[26]. The issue of ownership in the sense of

[26] It has been pointed out to me that this could be because in post-soviet systems there is a fear that anything that is not formally assigned to a single agency will not get done. It is believed that aversion to shared responsibility dates back to the late 1990s when many laws were passed which assigned responsibilities simultaneously to different levels of government. Resulting confusion led to significant problems. This was seen to allow buck-passing. Another criticism of that period was that the finance system and the system of distribution of responsi bilities seemed unconnected. Reformers wanted to fix a budget source to each responsibility so that there would

involvement, leadership and acceptance was never raised. It was the case that while compulsory indicators were "owned" by the federal government, the responsibility for monitoring and reporting them was assigned to regional or local governments.

Perhaps, this can be taken as a sign of certain underestimation of the role of "ownership" by the authors of the reform. This may indicate some insensitivity to the importance of the interplay of vested interests in the process of implementation of a reform. The importance of recognizing the seriousness of bureaucratic resistance was recognized already by Max Weber, who pointed out that, without the consent of civil servants, a monarch is unable to rule his state apparatus: «The Russian czar of the old regime was seldom able to accomplish permanently anything that displeased his bureaucracy and hurt the power interests of the bureaucrats» (Weber 1948: 238). Bureaucrats' role in successful reform is as great if not greater than that of "political will"[27] (Elmore 1980; Pressman and Wildavsky 1984). Within the tradition of New public governance consultation and consensus are seen as important instruments in overcoming this resistance and securing "buy-in" of the staff[28] (Barber 2007; Pollitt 2006a).

---

be no 'unfunded mandates'. This aversion to shared responsibility could be an obstacle for implementing PM today.

[27] It seems that the difference between Weber's notion of a "bureaucrat" and the modern notion of a "civil servant" is that Weber talked about relatively powerful elite among bureaucrats who could act as political subjects, whereas what is often discussed in modern managerial literature is the overcoming of resistance of the "rank and file" of civil servants. It may be argued that many modern reforms in the public sector, in fact, are generated by the bureaucratic elite and targeted at entrenching their domination over the lower levels of the bureaucracy and the population. Interests of front-line staff may be opposed to the interests of the managerial elites, while all of them together are called "bureaucrats"/"civil servants". It is interesting to consider how "political" is the "political will" today. It could be that it has, in fact, become the "bureaucratic will" masquerading as "political".

[28] From a critical perspective it may be argued that consultation procedures only mask the real nature of many reforms behind the façade of artificial consensus. The aim of many organizational reforms, both in the public and in the private sector, may ultimately be seen as the redistribution of power within the organization. The term "employee resistance" in the public sector literature is often used as euphemism meaning that employees are only resisting the reform because they are ignorant and short-sighted, whereas the designers and implementers are benevolent and wise. By going through the motions of the consultation process the employees are presumed to become more educated and more aware of the unity of their own interests and the interests of the organization. It may be argued, however, that 1) such process may also be interpreted as the process of ideological conditioning by means of which employees are forced to believe that the reform is in their own interests, whereas it really is not and 2) the consultation process may work as a "safety valve" for a harmless, moderated and "domesticated" expression of dissatisfaction and protest that is brought within the bounds of the system and thus dissolved. So that through the process of formal consultation antagonistic interests of employees are channeled into the harmless and meaningless activity of attending meetings, drafting reports and proposing changes. These proposals may then safely be dismissed or watered down to ensure that the original reform proceeds with only minor

The exploration of the process of implementation was an integral part of the original Pollitt's study (2006a). Within the frame of the present project respondents were asked whether they were involved in the process of choosing and formulating indicators or were involved in any other aspect of developing or modifying the system of performance measurement.

It was observed that, in terms of feedback, attitudes towards performance indicators differ depending on whether staff members were involved in choosing and setting the indicators. Two types of indicators may be identified: those that were imposed by federal authorities from the top and those that were generated from within the regional government. For the former type there was little feedback – the perception voiced by some respondents was that suggestions and comments sent to the federal government were often ignored:

> *The region is not involved in the process of selecting these indicators. We were not consulted while the methodology of performance evaluation was being framed. Neither when the DRONDs were introduced. [CS-1].*
>
> *The DRONDs came from the top, and that was it [CS-2].*

The process of selecting compulsory performance indicators for the federal list was seen by respondents from one of the regions as a mystery (in other instances these respondents also expressed concerns about the fact that their governor was not proactive enough and was excluded from the process of discussion of the system at the federal level). Federal compulsory indicators were considered as something invented "by someone on the top." There was a certain alienation from these indicators and the process of development of such indicators was treated with scepticism:

---

alterations, while creating the impression of "building ownership, trust and the spirit of mutual respect" between managers and employees. Employees may thus be lured into believing that they partake in the shaping of their organization as respected stakeholders, when, in reality, what they are partaking in is the construction of their own prison. These interpretations are made even more meaningful for Russia, where there are numerous cases of rampant violation of basic workplace ethics and rights of the employees by their superiors and where trade unions are either non-existent of powerless in the public sector.

They must have taken these indicators from Europe, changed them a bit to correspond with the Russian reality and rolled out. And that was it. Now we must work with it, whether we like it or not [CS-3][29].

Indicators that were internally generated were perceived somewhat differently. They were included in the annual Report on results and activities (DROND) together with mandatory federal indicators. Respondents expressed a more positive attitude towards these indicators. For instance, one indicator was viewed as an achievement:

We use QALY (Quality Adjusted Life Years - ASK). This is quite unique, other regions do not have such an indicator, but we use it in our rehabilitation centres. This indicator is among key measures in some of our programmes [CS-2].

It is also interesting to note the perspective on the feedback process of those at the other end of the managerial chain. I was fortunate to be able to interview the person who was responsible for managing the process of developing and implementing the system of performance measurement at the federal level in 2004-2007. This senior level federal civil servant stated that regional accounts about the lack of feedback are often over-dramatized and aim at avoiding blame:

Interviewer: In your opinion, was there a feedback mechanism in the process of implementing performance indicators into the practice of regional governments?

[29] It could be that this perception was a necessary result of the reform being crude when it was first tried, but this perception of the system as "foreign" and "imported" expressed by this respondent may, indeed, be a sign of a more general skepticism. If the reform fails to meet local needs and answer local demands it has no chance of taking root and evolving. If the reform is perceived as inauthentic it could remain an artificial supplement and would have little chance of having a substantial effect on the practice. Other data seem to correspond with this assertion. The civil servant from the federal Ministry of Finance who co-authored the initial design of the reform and then was involved in its implementation observed that regions were complying with the rules of performance measurement only until there was federal pressure. Once this pressure ceased, the practice of carefully cascading aims and linking them to indicators also tended to come to an end:

We were actively working with every ministry to explain how to link aims to tasks and targets to indicators. And we could see the progress. People really started thinking in these new terms. It was a huge improvement from what we used to have. They only used to think about dividing the money, but now they also had to think in terms of purposes of the spending. But, regrettably, when our commission stopped working in 2008, DRONDS began to fade and wordings of funding proposals and reports reverted to their dullness and simplicity.

Moreover, this practice of rolling out half-baked solution across the entire federation has been lamented by some respondents from among the academics. Instead of first launching a pilot project in one or several regions, it has been an often regrettable practice to implement underdeveloped reforms in the entire country.

Respondent: Feedback always exists. It may be expensive, covert, blotted, distorted, noisy, but it always exists. People go back and forth, meet each other, consult experts, do some research and so on. There is always some feedback.

Interviewer: When I interviewed civil servants in regional governments they complained that feedback was lacking and told me that their proposals were often ignored by the federal centre.

Respondent: Well, of course they did and rightly so. It's a tradition to complain. If you ask federal civil servants they would complain in the same way about the regions. Everyone complains about everybody else. There is a kernel of truth in it, but there is also enough slyness in it. Of course, if they wrote a useless document they then would complain that it was ignored. I believe that the rule is that "the reader is always right". If their document was ignored, it probably meant that it was a bad document.

Informal managerial functions of performance reporting

Despite the overall scepticism towards the system, the study has, in fact, identified management functions for which compulsory federal performance indicators were perceived to contribute. These functions were: 1) problem diagnosis, especially as a result of cross-regional comparison of performance data and the motivation created by the psychological effect of league tables, and 2) signalling function in attracting funds to critical issues.

Psychological effect of rankings as stimulus for competition

Respondents pointed out that the use of performance indicators for ranking regions in terms of government effectiveness created a psychological effect and encouraged governors to seek causes of poor performance of their regions. It is worth noting that none of the respondents considered federal grants an important incentive in the quest for better performance. Formalized rewards were considered insignificant[30]. The "psychological effect", on the other hand, was mentioned as an important factor. One respondent (deputy minister of economic development) explained this referring to highly competitive spirit of governors and their preference for leading positions:

[30] During the first two years of operation of the system 2bln rubles (~ $60 mln) were distributed in performance-based grants to 20 top regions. In following years 1bln rubles was distributed to 10 top regions. This amount of money is relatively insignificant if compared with regional budgets.

> *A competition of regions is a competition of leaders ... I have repeatedly witnessed that, on seeing his region's very low ranking, a governor would ask himself: "Why in this area?". And here the work would begin: "How can we fix it? Is it possible to deepen and expand and improve?'. Thus, I think that, despite all the criticism of the methodology, this ranking exercise has not only a psychological, but also an analytical benefit [CS-1].*

This emphasis on "psychological effect" and the competitive spirit of governors is consistent with resent research done on criteria of career development of Russian governors done by Yakovlev (Libman et al. 2012b; Yakovlev 2010). Yakovlev remarks that, not only for governors but also for other civil servants, criteria of career promotions are not formalized and are highly subjective and irrational. Since these are not formalized, governors may well consider it important to demonstrate leadership in performance rankings, even though this does not secure any significant formalized and predictable benefits for them. In a situation where career prospects depend on arbitrary decisions of the principal, agents may be expected to attempt to maximize their chances for promotion by demonstrating good performance even in substantially irrelevant exercises, "just in case"[31].

Highly arbitrary nature of career promotions in the Russian civil service has also been observed in a representative survey of civil servants by the Higher School of Economics[32].The survey found that over 73% of respondents thought that the most relevant criteria of career

[31] I believe that this particular kind of psychological effect may be associated with the fact that during the research period governors were appointed by the president. It was ambiguous whether a governor's position was a political or a purely bureaucratic one. Probably, it was some sort of a mix of both. For example, it was believed that governors were indirectly held accountable for low electoral support of the ruling party, even though officially this was never announced. Publication of performance data and league tables strengthened upward accountability, because the president could theoretically compare different regions on a single list of performance indicators. Thus, governors had to compete to demonstrate superior performance in the eyes of the principal, who could observe the entire federation "from above". In contrast, when governors came up for elections in 2013 and thus (at least theoretically) become accountable downwards, to the citizens of the region, it is unclear if the psychological effect of league tables would remain. After all, citizens have limited capacity to process data and have limited ability to travel to neighboring regions to compare conditions of life across the country. Even if citizens know abstractly that their region is doing badly compared to other regions, they may not be able to translate this abstract knowledge into practical political demands, or may be unable to "vote with their feet" and move to a more comfortable region because, for example, of transaction costs, such as costs associated with re-location and changing jobs. Thus, I believe that the observed competition-inducing effect was particularly pronounced because of the highly centralized and bureaucratized "vertical line of power". In a more devolved setting it might not be as prominent. This particular system may loose its relevance for elected governors.

[32] http://cinst.hse.ru/contract

promotion was the disposition of their immediate supervisor and a large proportion of respondents found criteria of career growth unclear and unpredictable[33].

It is interesting to compare the perception of regional civil servants with the opinion of the federal civil servant who was directly involved in designing and implementing the system. In his opinion this "psychological effect" was, indeed, among the most important tangible functions of the system, and other functions were mere "official rhetoric":

> *To be honest, I have never heard that anyone took this system seriously. Except, of course, for the official rhetoric about the fundamental managerial importance of the system. The only real reasonable function of this system is that it allows us to provide the president with an analytical brief when he is meeting a governor. So that the president can look at the table with performance figures and ask the governor "Why is the value of this indicator 5 in your region, while in other regions it is 8?". This is important, because it allows the discussion to begin. Then the governor has to respond and defend himself and so on. But functionally, there is nothing.*

He went on to reinforce his point:

> *All these rankings are, plainly speaking, nonsense.*[xiii] *In the end, they are symbolic. Of course, it is nice to get some grant, but really, people do not take this process seriously. This process is treated as just a tiny supplement (надстроечка). Nobody makes deliberate efforts to improve an indicator to get a grant. It requires a lot of efforts. These efforts are being made but the motivation is different. Grants are symbolic not motivational*[xiv].

According to him, league tables were important because they allowed regions to compare themselves against other regions, thus providing (often informal) benchmarks:

> *The psychological effect of league tables is more important. Through league tables regions can compare themselves to other regions across the country, so they can set benchmarks for themselves. They always compare themselves to others: they phone each other, consult and compare with competitors… They draw conclusions, and these conclusions need not be public or formalized*[xv].

[33] One of the questions read: "Who, do you think, is in the position to most accurately judge the results and quality of your work?". Responses: 73% immediate supervisor; 25% the head of department.
Another question: "How clear and predictable are the criteria of career growth in the civil service?" Responses: clear and predictable – 19%, all depends on the immediate supervisor – 24%, all depends on the head of department – 21%, criteria are unclear and unpredictable – 19%, do not know – 17%.

The emphasis on the symbolic function of indicators that allowed the president to have a bargaining advantage during negotiations is in line with another remark made by a consultant on controlling functions of indicators. He observed that the sheer number of indicators meant that they were a poor managerial tool but a good tool of control used by governors to supervise their subordinates:

> *Among three hundred indicators the governor can always find those that can be used to justify either rewarding a minister or dismissing him. [C]*

Among respondents there also were management consultants with extensive experience of working with and for the government. They saw performance indicators primary as "an instrument used by the federation to apply pressure on regional authorities". The sheer number of indicators, non-formalized sanctions associated with their growth or decline and the absence of clear prioritization between the indicators created a situation of high uncertainty. In any event the superior authority could “likewise either criticize or praise" a governor or a minister. It largely depended on political considerations. Thus, an image of complete rationality and objectivity was maintained amidst universal subjectivity and arbitrariness:

> *These indicators are mutually exclusive. If some are good, then some others must be failing. Indicators ... - are just a tool of pressure on regional authorities ... Among them one can always find an indicator for which someone is to blame. This is not a management tool. It is a tool of administrative pressure [C].*

All respondents commented that there were too many indicators and they were not harmonized. A consultant illustrated this problem giving an example from his experience:

> *...I reviewed all healthcare government programmes in the region: four federal target programmes and three regional programmes. I compared the number of indicators and their overlap. In four large programmes there were a total of 294 indicators, of which only 10 overlapped in at least two programmes. Three indicators were used in at least three programmes. And only one indicator – infant mortality - was common for four programmes. [C].*

This usage of performance indicators lies beyond the technocratic domain of public sector management. Perhaps, a better explanation may be given in terms of power struggles between interest groups[34].

A consultant summarized his impression of working with the system in following words:

> *These indicators are just a cudgel (дубина) to be held over the governors' head for discipline.*

Signalling function

Respondents identified another function of performance indicators as important. Indicators with negative dynamics were purposefully included in the DROND to influence budgeting decision and attract funding to the issue:

> *One of our [internally generated] indicators demonstrates negative dynamics. It's a bad indicator, but we have included it on purpose. It signals the need for more funds. We keep it to make sure that funding continues [CS-2].*

However, the same respondent noted that the relationship between performance data and funding was not formalized. She was at a loss to explain how it worked. It was possible, the respondent remarked, that the signals given by such indicators did not translate properly into budget decisions. But they hoped that somehow it did help them to get more funding through some informal means. It is, indeed, not clear whether signalling function worked in reality. For example, another respondent – a former regional minister of economic development - summarized the role of performance reports in budgeting process as follows:

[34] As noted by L. Yakobson, '(Russian) Budjetnaya Reforma: Federalism Ili Upravlenie Po Resultatam (Budget Reform: Federalism Vs Performance Management)', *Preprint WP8/2006/03, Gosudarstvennoe i municipalnoe upravlenie* (Moskva: Vyshaya Shkola Ekonomiki (Higher School of Economics), 2006)., an idealized system of performance indicators, would allow for a "perfect contract" between citizens and the government. It is, of course, not in the interests of politicians to approximate such a contract, for in a perfect contract accountability is also perfect. The current system, on the contrary, provides politicians with the means of monitoring bureaucracy, without creating a mechanism for the society to monitor politicians. Such a system of asymmetric control is, perhaps, not unique. U.S. researchers suggest that PART assessments could be used in a similar way V. Gueorguieva et al., 'The Program Assessment Rating Tool and the Government Performance and Results Act Evaluating Conflicts and Disconnections', *American Review of Public Administration,* 39/3 (May 2009), 225-45.. The interesting peculiarity of the Russian case may be the relatively higher importance of informal incentives and tacit rules associated with performance indicators. Formal procedures of rewarding regions for better performance were universally perceived as secondary to the informal system of incentives that spontaneously emerged and attached to the indicators.

I have a feeling that these are just documents for documents' sake. The DROND has nothing to do with budgeting. In this case, all its value is reduced to absolute zero [SG-1].

Any signalling has to take into account the expectations of the (potential) audience. Different stakeholders interpret the same performance data differently. As Behn (2005) observed, no indicator is usable for all purposes. In terms of the theory of speech acts, for example, the illocutionary force a performance figure depends on the recipient of the data. In the Russian case respondents did not seem to think about performance data in terms of different target audiences. They seemed to reckon that performance reports were mainly used by other governments. For example, the head of the directorate of capital investments was confident that businesses disregarded official statistics and used the word-of-mouth instead:

Interviewer: Do you think that businessmen use these data when they make decisions to invest in the region?

Respondent: They do, but they do not look at the indicators. They consult other businesses and consult their colleagues, who have more experience in the region. For example, if someone is a metallurgist and wants to know how easy it would be for him to obtain required permits in the region, he would consult other metallurgists. I mean that instead of checking official rankings they really prefer to use local knowledge of their colleagues. You can publish any ranking you want, but they will always check it against other businessmen' actual experience.

Another respondent was asked, whether regional parliamentarians were using the data in any way. She replied in the negative: "I cannot remember a single instance when we were contacted by our regional representatives to clarify or expand on any performance figure. I do not think they really use them at all". This may indicate that the only relevant target audience of the performance data are other civil servants (both superiors and "competitor" in other regions). Other target audiences seem not to be considered[35].

[35] Civil activists interviewed at one of public hearings complained that the indicators were bad and could not be used because the information was poorly structured.

Non-performance-based budgeting

Despite the officially declared aim of implementing "performance based budgeting" (*бюджетирование по результатам (БОР)*), the study shows that in practice budgets remain separated from performance reports. Studies in this area conducted by Russian researchers indicate that performance reports (DRONDs) were often drafted retroactively, to create an illusion of result-orientation:

> *In the absence of a clear procedure enabling agencies to link spending to targets, aims and programmes, public bodies continue to budget for their existing functions throughout the year. Then later, when they prepare the performance report, they retroactively try to fit activities that they funded into the framework of targets, aims and indicators. It is clear that such a method of "performance budgeting" is ineffective as means of planning*[xvi] *(Khabaev, 2010)*[36].

This study also concluded that emerging practice demonstrated the artificial nature of the "performance superstructure":

> *The following nuance should be noted: it has become common that the performance report is prepared by a narrow group of specialists of the headquarters of an agency. Most other staff, including heads of departments and sub-units are not involved in the process. In this situation performance report is not linked to actual day-to-day functioning of the agency. Hence, result-orientation remains purely formalistic (Khabaev, 2010)*[xvii].

This study was conducted in 2009-2010, so it captured the moment when performance reports (DRONDs) had been operating for over 5 years. Two years later, when the present study was carried out, respondents indicated that the practice of writing performance reports was fading away. This appears to be in line with the evaluation of the system as "imposed" and foreign. This may indicate that the practice of setting performance targets and aims did not manage to take root and now remains as a left-over because of administrative inertia.

[36] http://bujet.ru/article/83969.php

Even wider context: real priorities and decorative elements

The issue of managerial autonomy was discussed with one of the authors of the initial design of the federal performance management reform (formally a senior official at the Ministry of Finance). His responses point at the decorative function of performance management initiatives and hint that, despite official declarations the aim of "westernizing" public service delivery was never particularly high on the agenda of federal ministries:

> *Interviewer: How about the principles of increased local managerial autonomy? Were they implemented?*
>
> *Respondent: They were declared. And they continue to be declared, but the desire to control remains at the federal level. I think that this is not a merely rhetorical issue. This is a socio-economic issue… Then [in 2005] and now all the same words were being said about the need to diversify and devolve. But what we were really concerned with was how to ensure that our large businesses were promoted globally. This was the utmost priority. But it was as if we were really doing one thing but officially were declaring other aims. Of course, we know that it is important to live in the country, build roads, hospitals etc. So we were writing about this, but everyone implied that we were not writing about what we were really busy with*[xviii]*…*
>
> *So, since the real serious activity was behind the scenes, all the rest was more or less decorative*[xix]*.*[37]

This reference to "the second dimension" of public policy may help to resolve the apparent contradiction between declared principles of results-orientation and the observed tendency towards increased process-control.

To make a transition to subsequent research findings one further element remains to be covered. It concerns the issue of trust in performance data. How trustworthy are data if they are perceived to be collected for symbolic purposes and are relatively disconnected from budgeting or managerial planning? Does the observed scepticism towards performance measurement effect data reliability?

[37] It is worth noting that performance management was listed as the number one direction of reform efforts in the Concept of Administrative Reform (2005-2010) (Klimenko, 2012). It is as if the most symbolic and ornamental of elements should come first and other "down to earth" proposal should be safely placed somewhere behind it.

(Dis)trust in performance data

The issue of trust in performance data was selected for further investigation during the second half of the research project. Partly this was due to direct reference to data manipulation made by respondents and partly due to theoretical expectations. Presumably, if performance measurement system is treated with scepticism by social actors then it is likely that data quality is low.

Among direct references to data manipulation there were both serious examples from actual experience and half-fictional anecdotes. For example, a deputy minister of health gave the following ironic reply when he was asked if his ministry was able to influence its performance indicators:

> *Well, how is it possible to influence the birth rate? Can the Ministry of health influence the birth rate? We've had such an example. I tell this as a joke, but this was a true story. In 1974, when the Chief Physician of Russia (РСФСР) reported to the Minister of Health of the USSR, he was reproached because the birth rate was falling. He replied "I'll personally see to it!". And the next year birth rate grew. You see what I mean? Now we joke that if you give a personal order to a chief physician he can improve any indicator for you, even the birth rate.*[xx]

This ironic example hints at a serious problem. If a government agency is under pressure to demonstrate improvement of its indicators and is unable to achieve this by policy measures, it may resort to data manipulation. The same respondent continued:

> *The dynamics of indicators does not depend on managerial decisions. Except for one decision: if you need to lower the sickness rate, you can simply stop recording cases of this sickness and the indicator would fall. A given year's performance then would be excellent. The same may be done with mortality rate. For example, with certain techniques you can hide tuberculosis-related mortality under various diagnoses. Write off a couple of hundred people and mortality would fall. In this case, we can manage our performance indicators.*[xxi]

Another respondent described the early practice of reporting performance. According to him, when the system was first implemented (in the Ministry of culture) the issue of potential data manipulation was widely discussed. Civil servants, first of all, wanted to understand

whether reported figures were going to be audited, because this would limit the scope for potential manipulation:

> *Interviewer: Have these indicators been discussed within organization?*
>
> *Respondent: Not in the last couple of years. When performance reports (DRONDs) were first implemented, there were a lot of discussions. I used to work at the Ministry of Culture. It caused a real stir, because some of these indicators are impossible to validate, so they could be "painted" discretionally (нарисовать можно от балды). Everyone wanted to know if anyone was going to check. This was the first thing. Then later the commotion gradually faded (как-то замялось). It became clear that no one was going to check.*[xxii]

Examples of "painting" (inputting figures at will) were known to the respondent from his experience of collecting data from municipalities:

> *Some of these indicators cannot be checked. For example, the "share of population taking part in organized cultural events". There was such an indicator, it was impossible to check. I knew that our municipalities (районы) were painting it discretionally (рисовали)*[xxiii].

During the first years it was unclear whether any formal audit would operate. So, according to the respondent, a lot of efforts were put into the reports. Many deadlines were missed and not all data were reported promptly. Then the experience showed that no audit was going to operate and procedural difficulties stopped:

> *In 2008 there was a monitoring exercise to evaluate the practice of preparing performance reports. After it, when all went still, people realized that figures could be simply painted without leaving their offices (*можно рисовать, сидя у себя в деревне*). All went on very smoothly. They stopped missing deadlines and began reporting figures that we wanted from them (*В срок и с цифрами, какими надо*).*[xxiv]

Other indicators that theoretically could be audited seemed to cause some real trouble. There was an example of unintended consequences that effected real behaviour. Public libraries were required to increase the number of books in storage. As a result, they stopped writing off outdated volumes to be able to show positive dynamics of this indicator. This indicator could be checked, so they refrained from outright manipulation and instead distorted their actual practice:

Respondent: They started clinging to these ancient tomes. These Marxist and Leninist collections were kept to help them meet this performance target. The idea was that they would have to buy new books to meet the indicator, but they had no money for it, so instead they kept this old junk. But they had ideal performance.

Interviewer: Why did not they simply manipulate data?

Respondent: They were afraid to do it. This indicator could be checked, unlike the one with cultural events.[xxv]

One particular example should be noted. In this case the pressure to demonstrate improving performance led to systematic inflation of performance figures. The actual performance was worsening, but reported figures were showing consistent improvement:

When I worked at the Ministry of culture we had to monitor our municipalities on the number of social clubs in rural areas. This number had to grow. These social clubs were in "houses of culture" (дом культуры – leisure centre). In villages there are only public ones left. They are closing down because they have no money and it is generally pointless to continue funding them there. So, the number of social clubs was falling. So what could I do about it? I used to call all municipalities and tell them "Do you understand what you should do with your social clubs indicator?" (implying, that they should put in an appropriate figure). They would say "OK, we get it". So they would paint a proper figure.

*Then, after a while, we realized that year after year the gap between actual and reported performance was getting wider. So we wondered how long we were going to grow this indicator in this artificial way? Fortunately, soon the idea of performance reporting started fading away (*с ДРОНДами всё замялось*). It was a relief.*[xxvi]

Other respondents also mentioned that "it was unclear if these performance reports are going to continue", "DRONDs are fading away"). It is interesting to note that the respondent mentioned that municipalities stopped collecting data when the regional government stopped pressuring them. It only became evident when the new governor demanded an update on municipal performance:

Now we've got a new governor. He wants to get an analytical report on performance across the region. I have been calling to all our municipalities. It's been two years since I last requested an update. So, I asked them to send me this year's data. They replied that no one was collecting these data anymore. I wondered why this was the case. They said "Well, they've stopped forcing us to do performance reports, so we've dropped it"[xxvii].

Conclusion

The overall impression created by the bulk of qualitative evidence is that of widespread scepticism towards the system of performance measurement. It appears that at the moment the system mainly performs decorative functions. This does not mean that performance indicators do not affect actual life of government agencies. But their effects seem to be predominantly indirect and informal. Various interpretations of these effects have been recorded: some respondents emphasized the intelligence function of performance data that allowed for critical issues to be identified, other respondents highlighted the role of performance data as a catalyst of inter-regional competition between leaders, still others interpreted performance measurement schemes as means of bureaucratic and political control and centralization. It should be noted that in all these interpretations respondents placed informal stimuli and incentives first. Formal rewards and sanctions explicitly outlined by the legislator seem to play a rather insignificant role compared to the network of informal rewards. The system became an element of a wider bureaucratic machine and was moulded to fit into the existing network of relations of power, control and accountability. In Western literature, performance management initiatives are sometimes depicted as drivers of change in public sector mentality (Barber, 2010). Performance management is credited with playing an important role in inducing result-oriented culture. Our findings suggest that the task of changing public sector mentality was, indeed, to some extend achieved through the introduction of compulsory performance reporting in Russia. The respondent familiar with the process of design and implementation of the system at the federal level attached high significance to changes in bureaucratic culture, stimulated by the introduction of elements of performance management. He observed that departments began "thinking in terms of targets, aims and policy priorities". He has, however, also observed that this change was short-lived and its effects seemed to fade when federal pressure was withdrawn. Respondents from regional governments gave a similar account – both regions and municipalities were preparing

performance reports and linking activities to targets until this was demanded by the federal centre, but they tended to shed this burden once released from the federal attention.

At the federal level, at the same time, the attitude towards the performance measurement system seemed also rather sceptical. Frequent seemingly hasty re-designs, lack of genuine feedback, absence of a comprehensive audit created an impression of relative neglect. After more than 7 years of operation the system remains poorly integrated in the overall business of government agencies and continues to give an impression of an artificial supplement. Managerial studies by Russian researchers show that performance reports continue to be used to create an impression of results-orientation instead, whereas overall activity continues with little regard to principles of performance management.

The aims of moving towards genuine performance-based budgeting continue to be announced and performance indicators continue to be declared an important part of the modernization of the public sector. Respondents involved in the process of implementation of the system recognize that this capacity of performance management to create an air of "innovativeness" and "modernity" is among its chief advantages. It is convenient to continue declaring principles of performance management to signal that our public sector is using best practices.

Those respondents who considered the current system of performance indicators valuable mostly spoke about its symbolic value. It is regarded as an important instrument in rhetoric and bargaining between the president and governors because it may provide a bargaining advantage to federal authorities. Large number of indicators provides plenty of opportunities for blaming and shaming.

Formal rewards and sanctions assigned to performance indicators were considered negligible compared to informal reputational effects of league tables. "The competition of region is a competition of leaders" and, therefore, any ranking is guaranteed high attention from regional leaders.

It is crucial to keep in mind that the system was in operation during the period when governors of all regions were appointed by the president. Therefore, the system of performance measurement was incorporated in the mechanism of bureaucratic competition within the federal hierarchy. This hierarchy, owing largely to its Soviet and even Tsarist legacy, is held together by a network of informal, tacit rules, stimuli and incentives. Hence the importance of what respondents referred to as "psychological effect".

The term "psychological effect" was used to describe instances when governors (or lower level officials) made considerable efforts to improve their performance even though no tangible explicit reward was associated with such improvement. The term was used to talk about intensified competition between governors induced by league tables. It is particularly interesting to consider consequences of such "psychological effect".

Any system of measurement (however neutral its intention may be), when it becomes embedded in a dense network of informal incentives within a hierarchy, inevitably acquires certain motivational gravity. When rules of promotion are ambiguous and depend on the discretion of the immediate superior, it is rational to try to please this superior by means of performing well in all possible respects, including hollow bureaucratic exercises of the "match-aims-to-indicators" kind and the "grow-this-indicator-no-matter-what" kind. Thus, it may be concluded, that, despite the fact that the national performance measurement system seemed to be initially designed predominantly for "intelligence" purposes, it, nonetheless, generated distortionary effects characteristic of "target" type systems. In particular, it generated incentives for data manipulation.

Respondents clearly indicated that for them data manipulation was a common and mundane phenomenon. Indeed, it stands to reason to engage in data manipulation if one perceives performance indicators as largely decorative but sanctions and rewards for their improvement as tangible and directly felt.

In the following chapter I concentrate on analysing performance data generated by the national system over the period of 5 years. I find that certain patterns in performance data may be theoretically interpreted as (at least in part) consequences of deliberate data manipulation. I report findings of a survey of municipal civil servants suggesting that indicators identified as suspicious on the basis of their dynamics are associated with lower trust in performance data. The fact that data manipulation was tolerated also supports the assertion that performance measurement system in Russian regions may indeed have functioned as mainly a decorative element.

Nonetheless, this decorative element provides valuable data for analysing bureaucratic structures. Even if performance indicators poorly reflect objective qualities of public services, they, nonetheless, reflect inner workings of the "vertical line of power". Passing through levels of hierarchy performance figures are distorted in a somewhat similar way to beams of light that are distorted by gravitational fields of stars and invisible black holes. Performance data may, thus, be studied to understand the invisible reality inside the bureaucratic universe.

The following characteristics of current performance management system have been demonstrated to limit the usefulness of performance indicators for managers: 1) agencies have little influence over performance measures that are used to evaluate their performance; 2) agencies demonstrate acute lack of interagency coordination and cooperation that prevents them from resolving issues that require input from more than one agency; indicators are not used to facilitate cooperation; 3) regional public sector managers perceive the lack of feedback in the process of selecting the set of indicators that is used to monitor their performance; indicators are imposed "from the top".

Performance indicators may potentially fulfil a wide range of managerial functions, but in the Russian case many of the functions remain unfulfilled. The table below provides a summary of the extent to which PM system fulfils its potential functions. Eight potential functions are taken from Behn (2003)

Table 13. Perceived managerial benefits of performance indicators in the Russian case. The list of functions adapted from Behn (2003).

Managerial function	Meaning	Perceived managerial benefit	Reason
Evaluate	How well is my public agency performing?	Marginal	Agency's contribution to the dynamics of indicators is small: external effects, long-termism, lack of cooperation
Control	How can I ensure that my subordinates are doing the right thing?	Marginal	
Budget	On what programmes, people, or projects should my agency spend the public's money?	Marginal	Determining role of federal decisions. Compulsory funding of federal mandates. Little local discretion.
Motivate	How can I motivate line staff, middle managers, non-profit and for-profit collaborators, stakeholders, and citizens to do the things necessary to improve performance?	Marginal	Alienation. Indicators are handed down from the top. Lack of involvement, communication, feedback and local initiative.
Promote	How can I convince political superiors, legislators, stakeholders, journalists, and citizens that my agency is doing a good job?	Marginal managerial benefit. Substantial symbolic benefits.	This may be regarded as a political (symbolic) benefit. Indicators create the veil of technocratic rationality.
Celebrate	What accomplishments are worthy of the important organizational ritual of celebrating success?	Marginal	Agency's contribution to the dynamics of indicators is small: external effects, long-termism, lack of cooperation
Learn	Why is what working or not working?	Some benefit identified	Cross-regional comparison may be useful for problem diagnosis, but lack of continuity in performance management systems undermines potential benefits for learning over time.
Improve	What exactly should who do differently to improve performance?	Marginal	Lack of interagency cooperation. Fragmented and rigid responsibility.

In theory, the introduction of performance management in the public sector can lead to institutional changes. However, in practice the opposite is often the case: "the overall institutional context is not modified and in fact shapes the way that the reforms are implemented" (Talbot 2008: 12).

CHAPTER 6. PERFORMANCE INDICATORS IN REGIONAL GOVERNMENTS: QUANTITATIVE FINDINGS

> *"In short, funny things are likely to happen when the bureaucratic strain of transparency meets blame-avoidance" (Hood 2007a: 208)*

This chapter presents quantitative findings. The analysis aimed at finding evidence of deliberate data manipulation. Qualitative findings presented in the previous chapter showed general scepticism of respondents towards the usefulness of the system of performance measurement. Some respondents reported cases of deliberate data manipulation.

Research efforts were directed at understanding common strategies of data manipulation and obtaining quantitative evidence of their effects. Since the database included performance data from all regions, quantitative analysis allowed for a statistical generalization from the three regions in the initial selection to the entire population of 83 regions.

Respondents were asked to characterize data manipulation strategies that they commonly witnessed in their practice of dealing with performance indicators. On the basis of these accounts two models were designed. These models allowed for hypotheses to be formulated and tested.

Once indicators exhibiting suspicious patterns were identified an attempt was made to estimate the trust in performance data generated by these indicators. The level of trust in performance data was measured by means of an online survey. A link between suspicious data patterns and trustworthiness of individual indicators was sought.

The chapter begins by laying out descriptions of two strategies of data manipulation given by respondents, then the analytical framework based on the principal-agent theory is described and hypotheses are formulated; then major pieces of quantitative evidence are presented that allow the hypotheses to be tested; then further findings that hint at wider research opportunities are outlined; finally, the results of a follow-up validation survey are presented that attempt to establish a link between identified oddities in performance data and the level of trust in data among civil servants.

Strategies of data manipulation

Two strategies of data manipulation were identified based on respondents' accounts. The first strategy may be called the "prudent bureaucrat" strategy. Civil servants using this strategy preferred to play safe. They reported only minor variations in the data as this was least likely to cause suspicion and attract attention of their superiors. No one could be blamed for a drop of 1% in annual performance, and a 1% annual growth would also be inconspicuous. At the same time, reported figures had to show *some* variation to look more "normal".

One management consultant and former head of performance measurement unit in a regional government described a situation where the "real" performance was not simply different from reported performance, it was unknowable:

> *When we find ourselves in the situation in which we don't have enough money to build or reconstruct roads it is unreasonable to expect from us to conduct surveys to measure precisely the share of roads that are in good or bad condition. So, if we are required to show the share of roads that are in disrepair we just input a number of, say, 67%. Then we give a forecast that next year it will be 66%, and the last year it was 68% because there should be some improvement over time. But no one can ever check this. I would say that about 10% of indicators are simply taken out of thin air (former civil servant, consultant)*

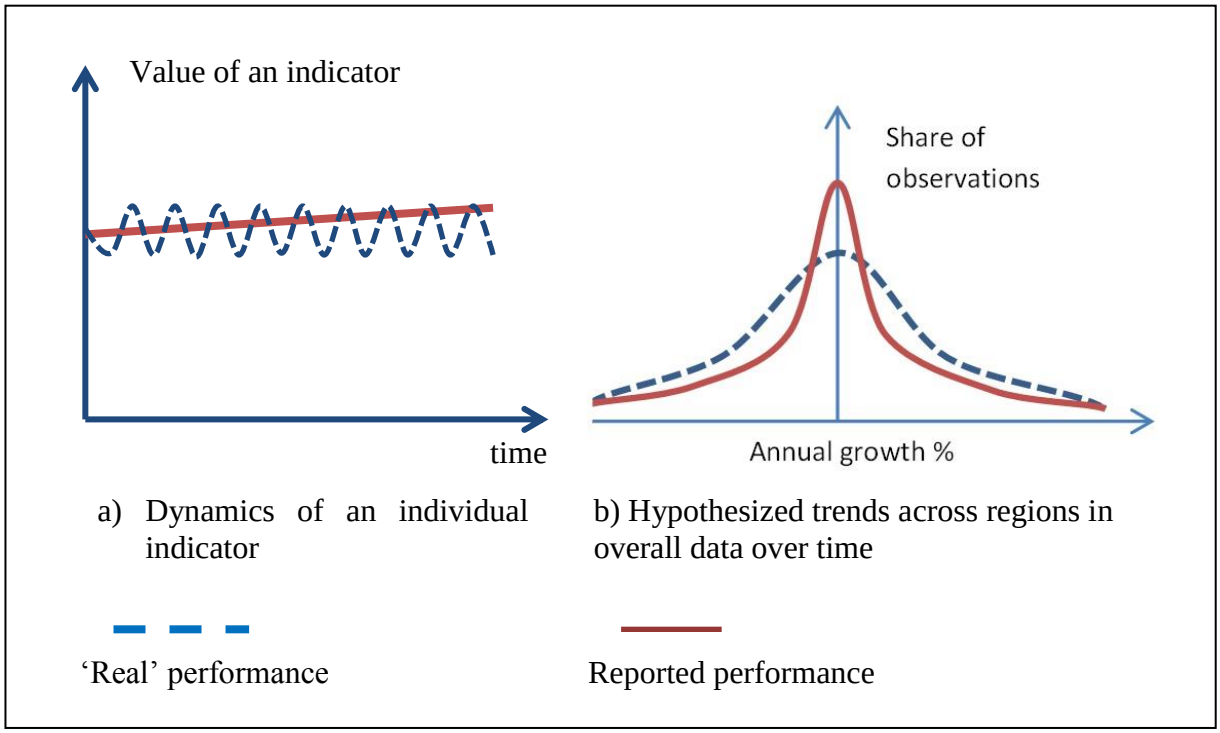
He considered this prudent method to be superior to a strategy of reckless inflation used by less far-sighted individuals. When asked if he thought that performance figures could be systematically inflated, he replied:

> *Well, I think that such inflation may only be done by very unwise and short-sighted men. I think that values of these indicators are reported in accordance with the responsible person's political sense and reasoning. You should not be too thin and should not be too thick, you should take the prudent approach and aim for the middle. (former civil servant, consultant)*

The graphs below demonstrates how "prudent" manipulation could affect the dynamics of an individual indicator and trends in overall statistics. If pursued systematically by a significant percentage of reporting authorities, this strategy would force values of growth indices of affected indicators to converge to zero in a "more-normal-than-real" fashion. This strategy allows the

manipulator to minimize the gap between reported performance and real life without giving an immediate impression of stagnation from year to year. Values do change from year to year, but in the long-run the indicator is kept stagnant. Allowing the gap to grow large is imprudent as it may result in severe sanctions if it is found out.

Figure 12. Consequences of systematic application of "prudent" manipulation a) for an individual indicator, b) for a set of indicators.



It should be noted that the "real" performance indicated by a dashed line is purely a hypothetical conjuncture serving the purpose of illustration. There is no such thing as completely objective data. All data are collected by someone who may be in a position to distort it. Moreover, any measurement is theory-laden and, in this sense, already not "real". Illustrations below attempt to show the expected direction and type of the bias introduced by systematic application of a data manipulation strategy. They are not intended to be making a statement regarding the nature of data before manipulation. For example a bell-shaped curve at figure 1b indicates that in the absence of deliberate manipulation growth indices would tend to approximate a bell-shaped distribution with a maximum at or around zero. It is not asserted that

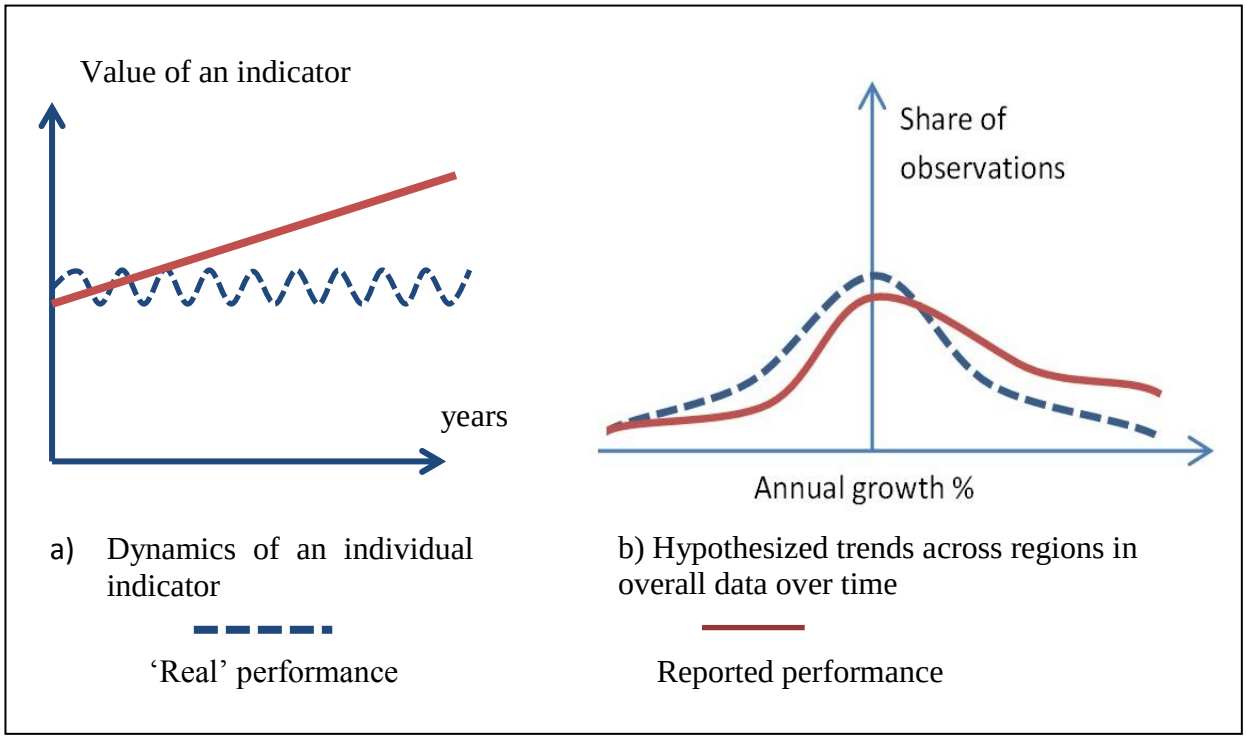
such distribution would be normal in strict mathematical sense, but it would be less concentrated in the proximity of zero than "prudently" manipulated data.

The second strategy consisted in systematic inflation of figures to report performance in a favourable light. One lower level civil servant gave the following specific example:

> *When I worked in the Ministry of culture, we had to show that we have a growing number of social clubs in rural areas. But there was no money and the people were leaving. So, actually, the number of social clubs was going down. But we had to demonstrate growth and we did. Year after year*[38]*. Then after several years we were relieved to find that we had no longer to report these inflated figures because the whole system was discontinued.*

If pursued consistently by a large number of reporting agents, this strategy would result in continuous inflation of performance data over time. Overall statistics would be biased in a way should below (Figure 13).

Figure 13. Consequences of systematic application of "reckless" manipulation a) for individual indicator, b) for a set of indicators



[38] In this example it is unclear what the motive behind continuous inflation was. The respondent's immediate superiors could be interested in either avoiding blame or maximizing credit. In any case the resulting statistics was systematically inflated.

The intensity of data manipulation of both kinds would depend on the model of principal-agent relations between the reporting body and the agent. Three such models were identified: A) impartial reporting agency; B) interested reporting agency; C) self-reporting agent.

Principal-agent relations in data reporting

Reporting agencies were grouped into four groups according to a model of principal-agent relationship that most accurately described their interaction in relation to data collection, reporting and evaluation. Two agencies were categorized as *impartial reporting agencies*. In the Russian case this model fitted the cases in which data were collected by either the Federal Bureau of Statistics (Rosstat) or the Federal Security Service (FSO). Neither of these two bodies was evaluated on the grounds of performance data that they provided. Federal statisticians could be considered relatively impartial[39]. Similarly, federal security agents were deemed to be above the game[40]. In this model there was assumed to be little or no incentive for deliberate data manipulation (model A. in Figure 14). Together these agencies provided data for 83 indicators (25%).

The majority of indicators (170 (53%)) were reported by federal line ministries (such as the Ministry of Education). The reporting entity was at the same time responsible for implementing government policy and could, presumably, be indirectly assessed on the basis of data that it collected. Thus, a conflict of interest could occur, but the possibility for manipulation in this case

[39] A recent confirmation of this may be found in the fact that the Federal Statistics Bureau (Rosstat) recently published population estimates that contradicted political statements of Prime Minister Medvedev. The politician announced that in 2012 the population of Russia demonstrated positive growth for the first time in decades and credited the government for this achievement. A few days later Rosstat published the results of annual survey that showed a minor drop in total population. This contradiction was commented on as a sign of political impartiality of federal statisticians. (Gontmacher, 2013, http://echo.msk.ru/blog/gontmaher/1003510-echo/).

[40] The Federal Security Service (FSO) is charged with collecting citizen satisfaction data. One civil servant commented on this fact: "*When we used to collect citizen satisfaction figures in the Department of Health, we used to get 90% satisfaction all the time. Once FSO took over, they started showing 35%".*

seemed relative low[41]. All federal line ministries were categorized as *indirectly interested reporting agencies* (model B. in Figure 14)

The data for remaining indicators (72 (22%)) were produced by regional governments themselves. Some data were collected by municipal units, but the responsibility for reporting the data to the central government was assigned to regional officials.

No provisions for formal central audit of these data existed. Most interviewed regional civil servants thought that "there must have been somebody who checked the data", but were unable to give details[42]. This setting was associated with the largest opportunity for data manipulation (model C. in Figure 14).

According to the most appropriate model of principal-agent relationship reporting were categorised into four[43] groups: 1) Federal Bureau of Statistics; 2) Federal Security Service; 3) Federal Line Ministries; 4) Regional Administrations. The total number of indicators that were reported by each group is given in Figure 14.

[41] If, for instance, the Ministry of Education were to engage in deliberate distortion of performance figures it would have to attribute higher ratings to some regions and lower ratings to other regions and such discretionary changes could cause dissatisfaction with regional administrations. It seems unlikely that such manipulation could be widespread, but it was not deliberately investigated in this research project.

[42] One regional minister of economic development was clear that he did not trust locally generated data: "However well qualified local civil servants are, one should never rely on their data. The growth of any indicator depends on the person who seats in the courtyard of his collective farm and counts. One should only rely on the data of the Federal Bureau of Statistics".

[43] I categorize the Federal Security Service (FSO) into a group of its own due to the peculiar nature of indicators that were assigned to this body. FSO collected citizen satisfaction data for all areas of governance (total of 8 indicators). The fact that satisfaction indicators were assigned to this organization may be interpreted as evidence of the low level of confidence in data provided by line ministries and regional administrations.

Figure 14. Principal-agent relations in data reporting

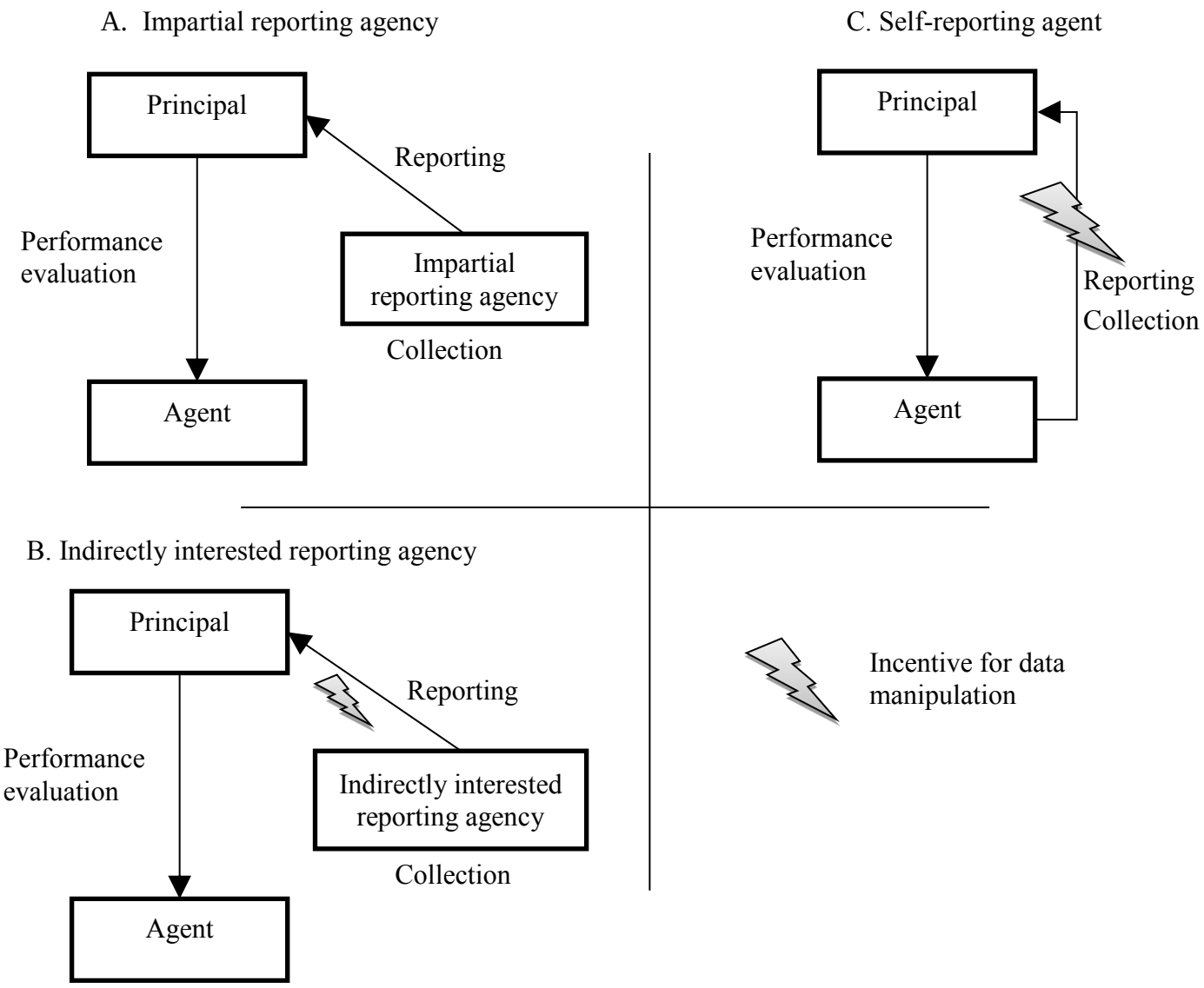


Table 14. Number of indicators in groups of agencies.

Group	Corresponding model of principal-agent relationship	N of indicators (% of total)	N of indicators used in ranking for grants (% total)
Regional governments	Model C – self-reported	72 (22%)	4 (6%)
Federal Bureau of Statistics (Rosstat)	Model A – independent agency	75 (23%)	14 (22%)
Federal line ministries	Model B – interested agency	170 (53%)	41 (65%)
Federal Security Service (FSO)	Model A – independent agency	8 (2%)	4 (6%)
Total		325 (100%)	63 (100%)

Hypotheses

Hypotheses regarding the distribution of annual growth indices are formulated based on the assumption that systematic manipulation of the data along the levels of bureaucratic hierarchy is detectable through observing distortions in overall resulting figures.

Figures 12 and 13 show the hypothesised difference between biased and unbiased indicators. Blue dashed lines represent the "normal distribution" of annual growth indices in a hypothetical situation where the data are collected in an honest and impartial way and the overall "real life performance" of the entire country is stable (there is neither an upward or a downward trend, the country on average is "treading the water", so to speak). In such a situation annual growth rates would be distributed in a way bearing resemblance with the normal bell-shaped distribution. Most regions would demonstrate insignificant growth/reduction with a minor share of outliers on both sides. If in such a stagnant situation, there is a deliberate effort to avoid reporting negative figures, or to demonstrate inconspicuous performance, the distribution would be skewed. The interpretation of resulting curves is similar to techniques used to capture effects

of electoral fraud (see, for example, Enikolopov et al. (2012). If humans deliberately interfere with the data along the data production process without grand coordination, a particular trend appears in resulting statistics that is not envisaged by the manipulators. In case of prudent manipulation the effort to be inconspicuous, if taken by a large enough proportion of reporting agents, produces effects that become themselves conspicuous.

It is hypothesised that both types of manipulation are more pronounced among self-reported indicators as there are both the greatest stimulus and the greatest opportunity to manipulate. The working hypotheses are:

Hypothesis 1: Prudent manipulation: Self-reported indicators are more stable over time. The distribution of growth indices converges to zero.

Hypothesis 2: Reckless manipulation: Self-reported indicators demonstrate more positive dynamics over time. The right-hand tail of the distribution is greater.

Hypothesis 3: Effect of grants: Indicators tied to federal grants are more significantly manipulated in the "reckless" way.

Hypothesis 4: Effect of grants: Indicators not tied to grants are more significantly manipulated in the "prudent" way.

Results – quantitative evidence of data manipulation

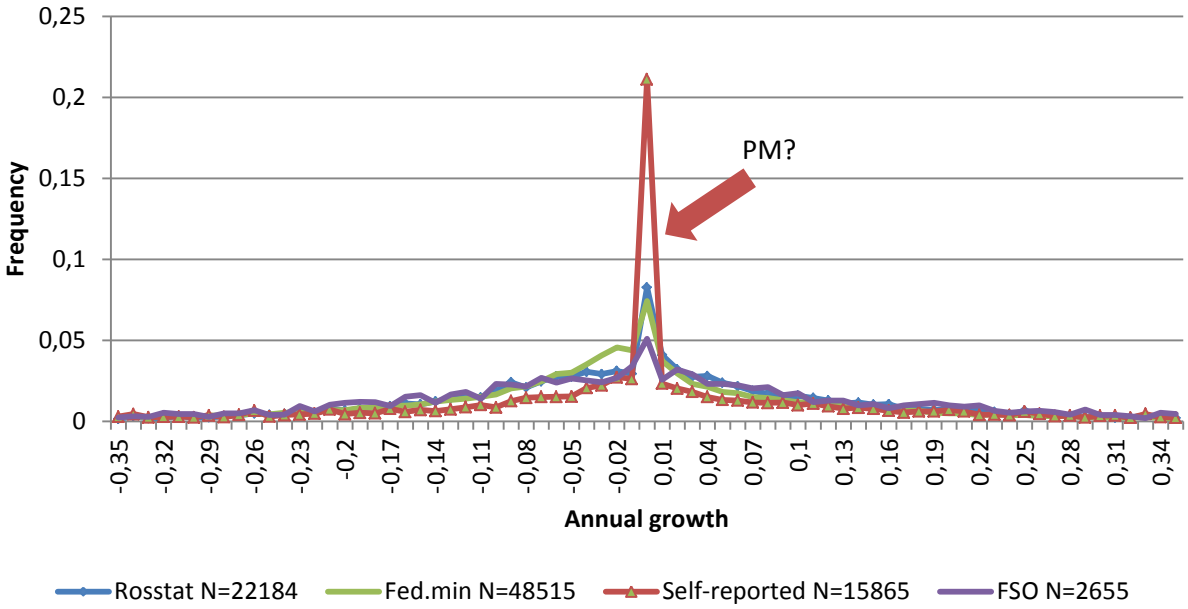
Prudent manipulation

The analysis of annual growth indices for indicators from the four groups demonstrates that self-reported indicators exhibit significantly different behaviour from the three other groups. In line with the "prudent bureaucrat" hypothesis self-reported indicators are clustered in the vicinity of zero to a much higher extent. This suggests that regionals officials could, in fact, adopt the "prudent bureaucrat" approach and underreport large variations in data. Significantly larger proportion of observations in this group demonstrates near-zero growth (Figure 15). The

symptoms of "reckless manipulation" are not visible in pooled data. This may suggest that, in general, prudent behaviour prevails among self-reported indicators.

Figure 15. Distribution of annual growth indices of indicators in four groups of agencies over the period of 5 years. [44]

Self-reported indicators (red line) converge to zero (red arrow) as predicted by the "prudent bureaucrat" hypothesis.



[44] Growth indices are calculated using the following formula: $\frac{PI_t - PI_{t-1}}{PI_{t-1}}$, where PI_t and PI_{t-1} – values of a performance indicator in years *t* and *t-1*, respectively. The lines demonstrate frequencies of occurrence of a given value as share of the total number of observations in this group.

Evidence of "prudent" manipulation

Figure 16 Distributions of growth indices of performance indicators for four groups of agencies in 2007/2008 period.

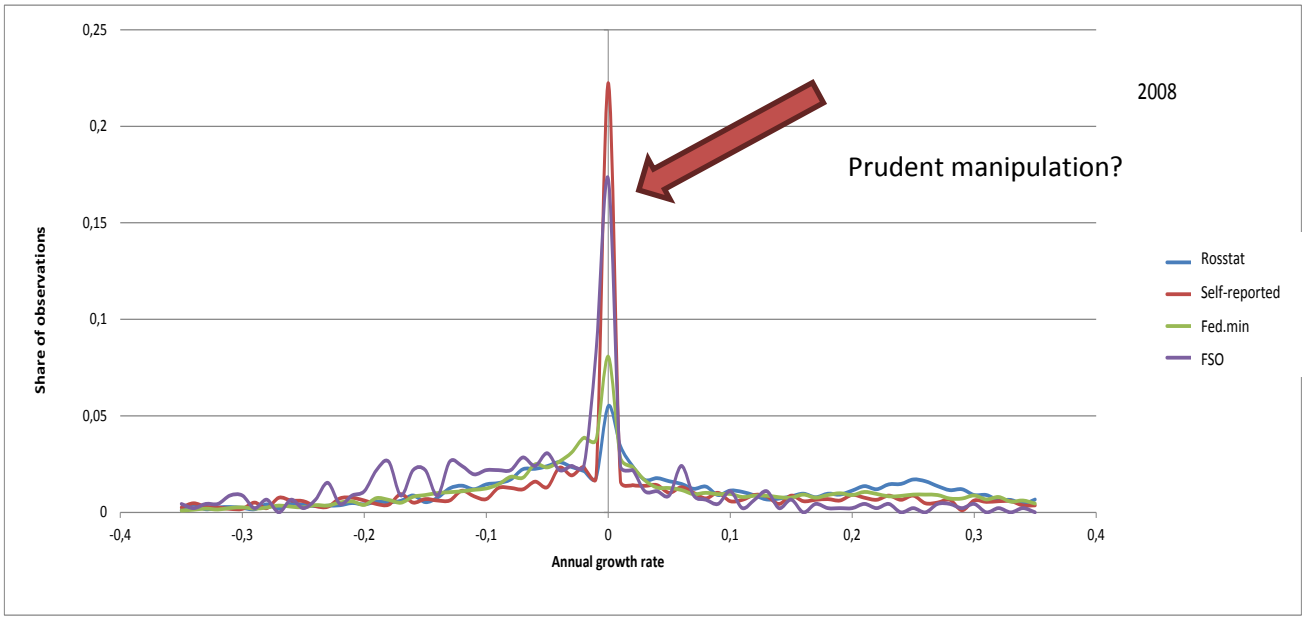


Figure 17 Distributions of growth indices of performance indicators for four groups of agencies in 2008/2009 period.

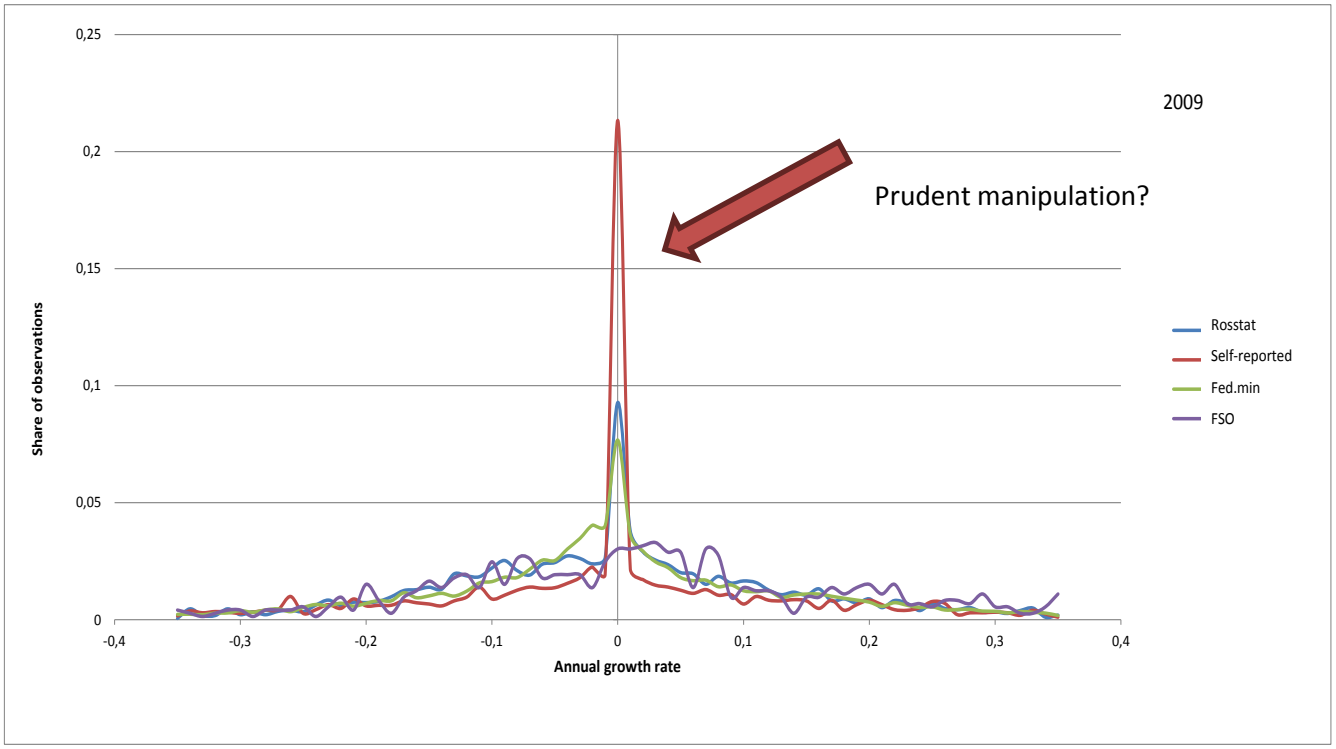


Figure 18 Distributions of growth indices of performance indicators for four groups of agencies in 2009/2010 period.

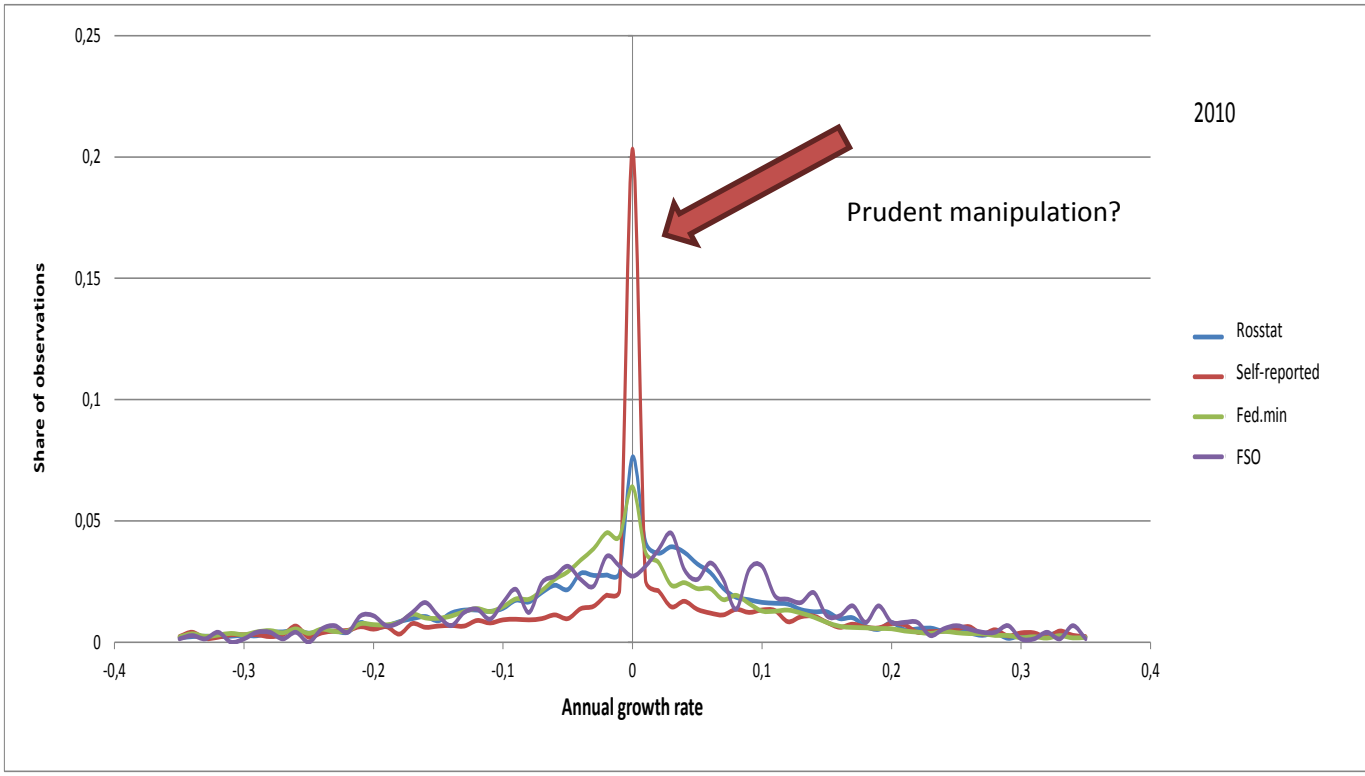
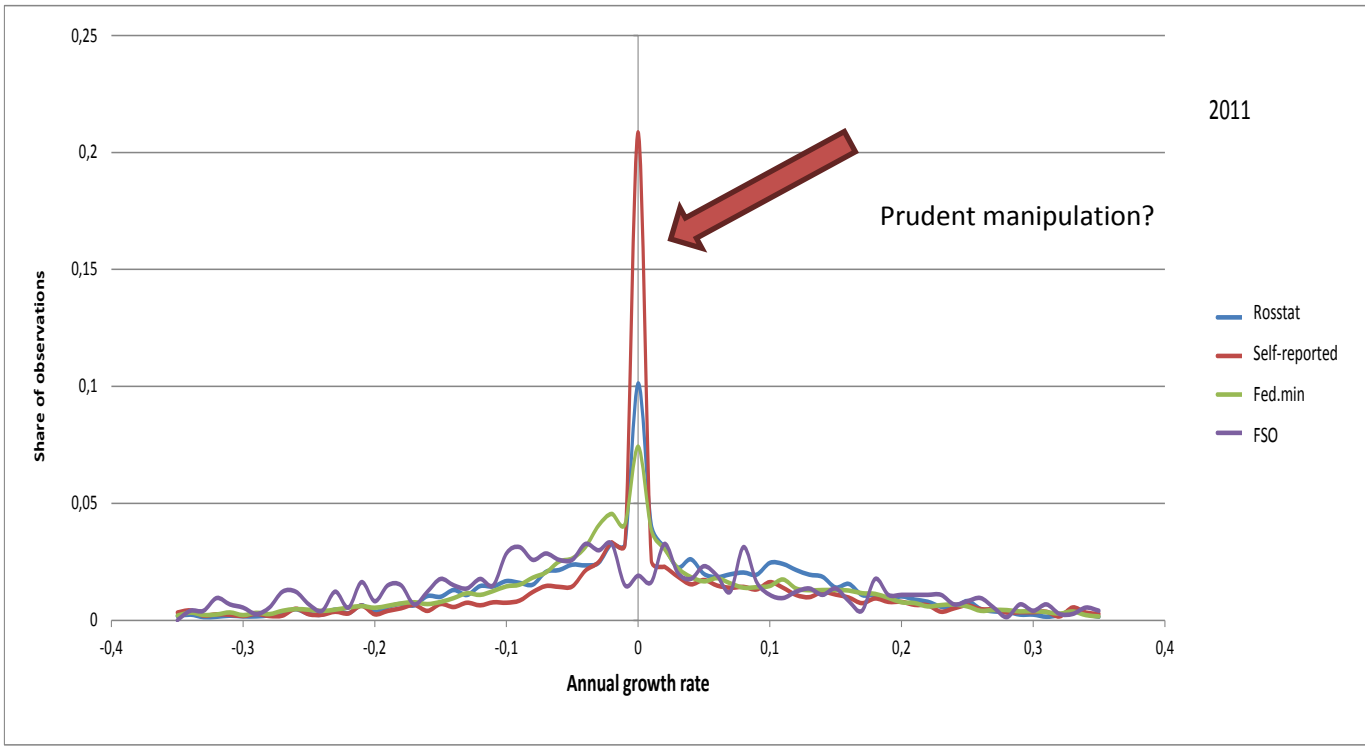


Figure 19 Distributions of growth indices of performance indicators for four groups of agencies in 2010/2011 period.



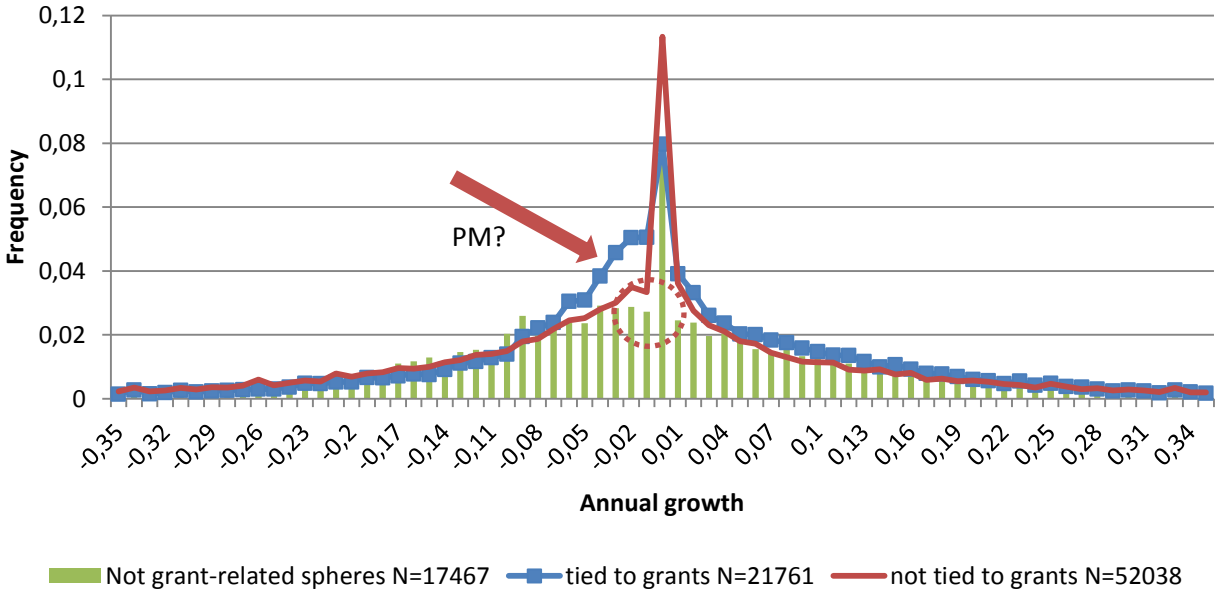
Effect of grants

It was hypothesed that indicators tied to federal grants should demonstrate significantly better dynamics. Additional perks for positive dynamics could stimulate governors to pay attention to values of these indicators. This could lead to both increased real activity to improve performance or create greater incentives to inflate performance figures. Figures below explore the effect of grants in greater detail. For this purpose indicators were broken down in two groups those tied and not tied to federal grants. Arrows in figures below highlight irregularities and kinks in distributions that may be associated with deliberate manipulation.

There were in total 224 indicators with positive polarity (those for which increase meant improvement, for example, birth rate) and 91 indicator with negative polarity (those for which decrease meant improvement, for example, mortality rate). 14 indicators were either "yes/no" indicators or had a specific target value set as a normative value. Cost indicators were corrected for inflation. The scale for indicators with negative polarity was inverted (growth rates were multiplied by (-1) so that they could be visualized together with indicators with positive polarity.

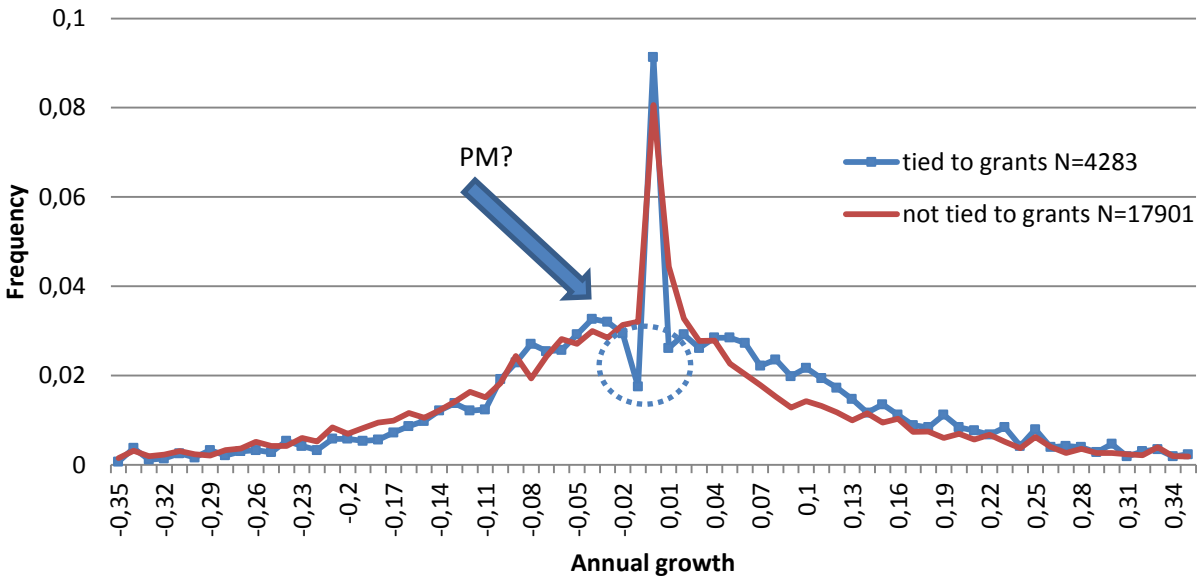
First let us observe the overall picture. Figure 20 below shows that traces of prudent data manipulation are present among indicators not tied to federal grants. They are even more significant among indicators from not grant-related spheres of public policy (such as vocational training and education, Table 2). This is indicated by negative kinks in the area of minor negative values. A possible explanation is that in the absence of clear targets or pressure from league tables civil servants adopt the prudent manipulations strategy to minimize labour costs of reporting and maximize convenience. Indicators tied to federal grants seem to demonstrate a more natural dynamics with no "kinks" in the distribution.

Figure 20. Evidence of prudent data manipulation (PM) among indicators not tied to federal grants. All agencies.



When the data are further broken down into groups of indicators reported by different agencies a different picture may be observed. Somewhat surprisingly, it turned out that, among the indicators collected by the Federal Bureau of Statistics, the effects of both prudent and reckless manipulation strategies may be observed (Figure 21 below). The negative kink in the left-hand tail may be interpreted as evidence of the reluctance of reporting agencies to demonstrate minor negative figures in indicators that were tied to grants. This suggests that some interference with the data occurs even within the system of federal statisticians.

Figure 21. Effect of grants. Traces of prudent manipulation (PM) among grant-related indicators. Rosstat. Pooled data for five years.

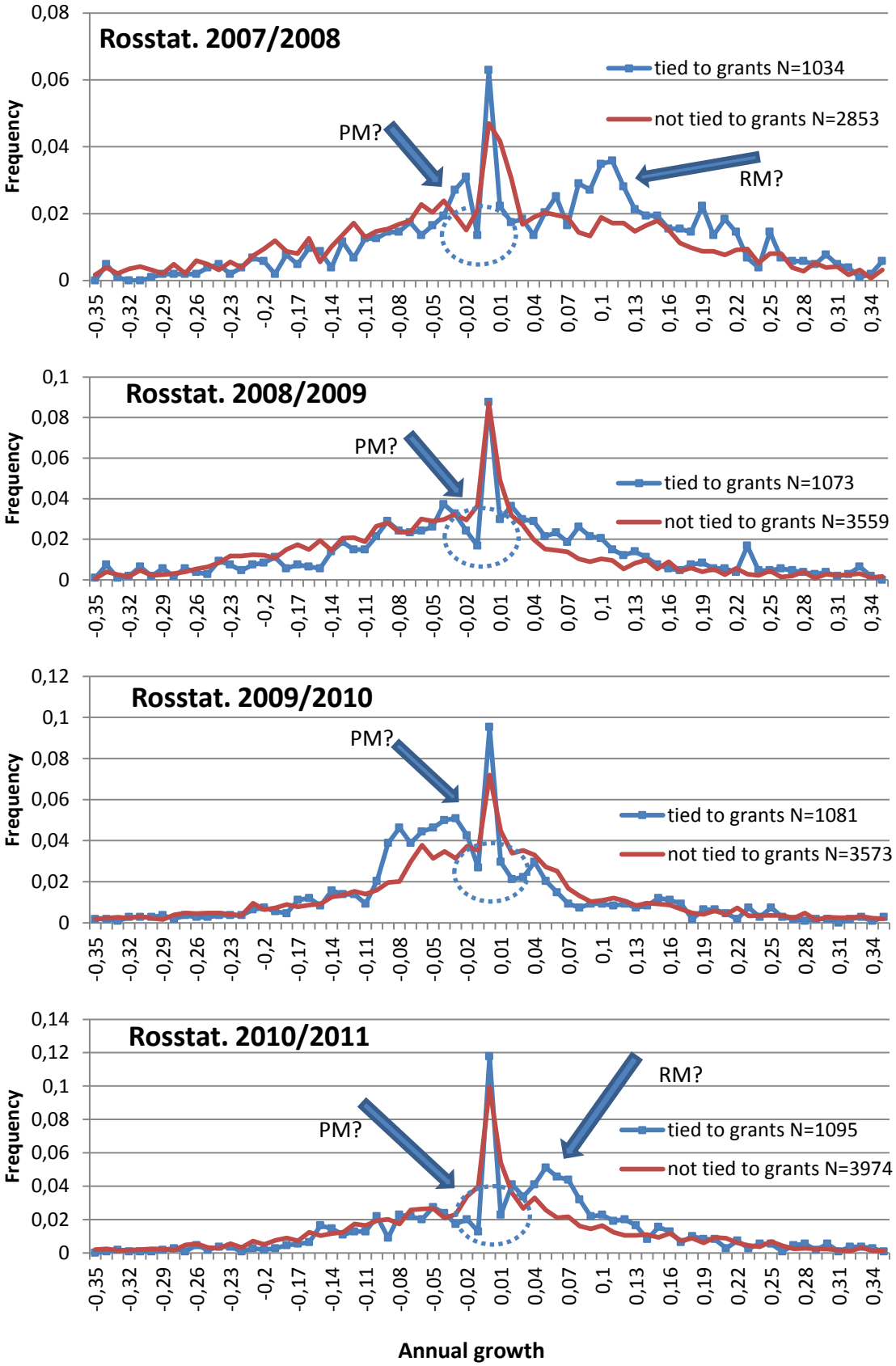


The collection of data may, in fact, be performed by the same people for different reporting agencies. The difference, however, is in the number of layers of "interested bureaucrats" that the data have to go through before they appear in the final report. Front-line data collectors may not be in the position to collude and cooperate to systematically distort the data in one pre-meditated way. But once the raw data are collected they have to be processed and prepared for publication. There may be numerous revisions by different people and the resulting distortion depends on the propensity of each person to "massage" the figures. It may be that the data reported by Federal Bureau of Statistics are subject to such distortion at some point in the process.

Another surprising fact is that Rosstat's indicators tied to grants seem to be affected by reckless manipulation (Figure 22). An alternative interpretation is that these indicators demonstrate genuine improvement from year to year. However, qualitative evidence speaks against such interpretation (See the section on (Dis)trust in performance data in previous chapter). Respondents remarked that indicators that they considered important could not be honestly improved in the short-run (see the section on Long-termism and External factors in the previous chapter).

Indicators reported by Rosstat

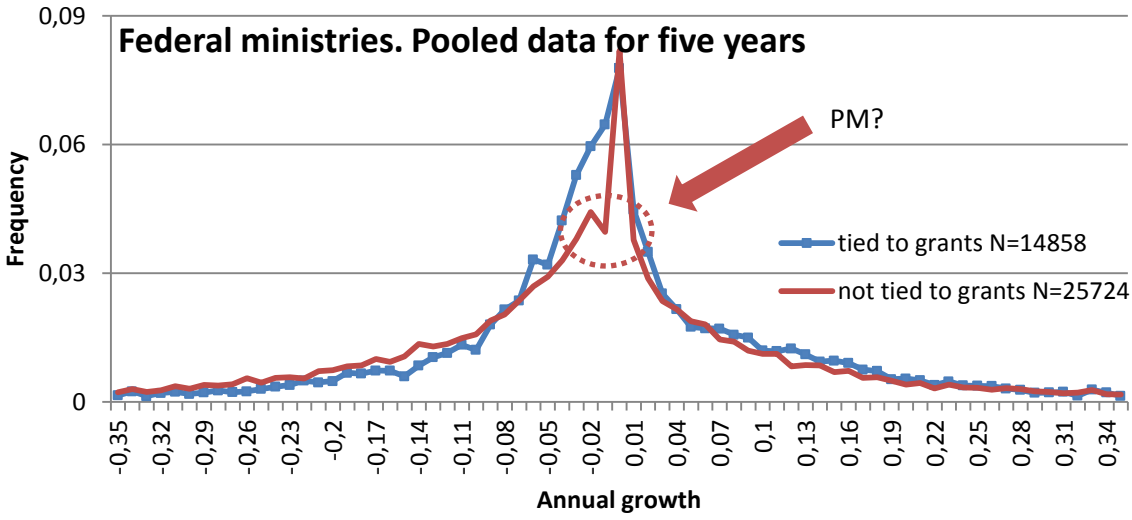
Figure 22. The effect of grants on indicators reported by Rosstat. Disaggregated data for each year 2007-2011. Arrows show traces of prudent (PM) and reckless manipulation (RM).



Note the significant upward kink in the distribution in the first time period (2007/08). This could be interpreted as an effect of publication of league tables. In 2007 suddenly the ability to benchmark oneself against other municipalities was introduced into the system. Poor performers could use this additional data to adjust their performance against the national average in 2008. The data were published by the Ministry of regional development and became available to the public and to regional administrations at the same time.

Let us now consider indicators reported by federal ministries. They present a different picture. Contrary to the expectation they do not seem to be affected by reckless manipulation. However, traces of prudent manipulation are observable in both pooled and disaggregated data.

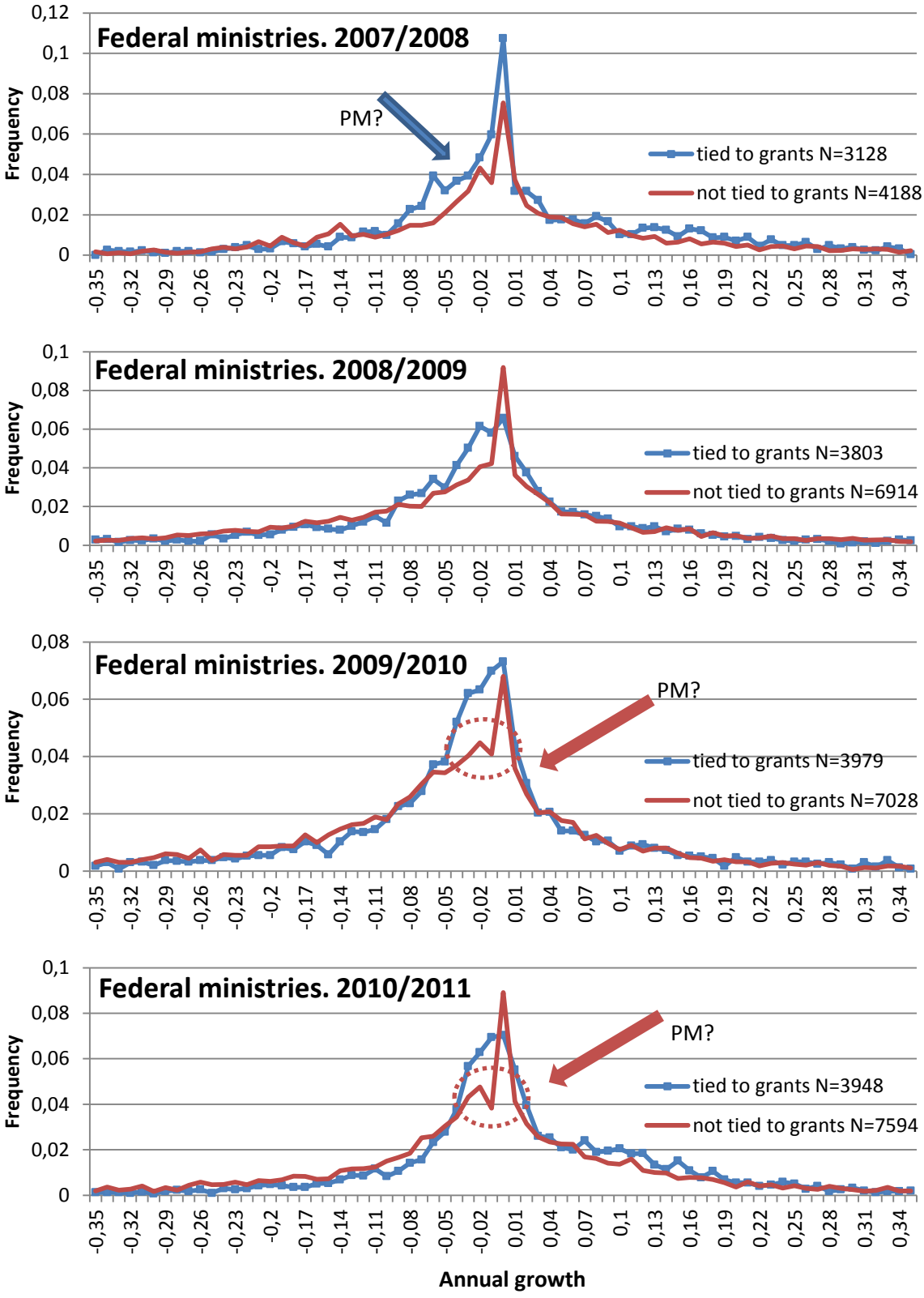
Figure 23. Effect of grants. Traces of prudent data manipulation (PM) among indicators not tied to federal grants.



The distribution of growth indices of indicators not tied to grants show a negative kink in minor negative values (Figure 23), whereas the distribution for grant-related indicators is more natural. Similar picture is observed in disaggregated data below (Figure 24). Except for 2007/2008 the distribution of grant related indicators is more natural than the distribution of not-grant-related indicators. This may suggest that grant-related indicators received more attention and data for them were actually collected at the individual level. In contrast, the data for at least some indicators not tied to grants could be input discretionally without actual data collection.

Indicators reported by federal ministries.

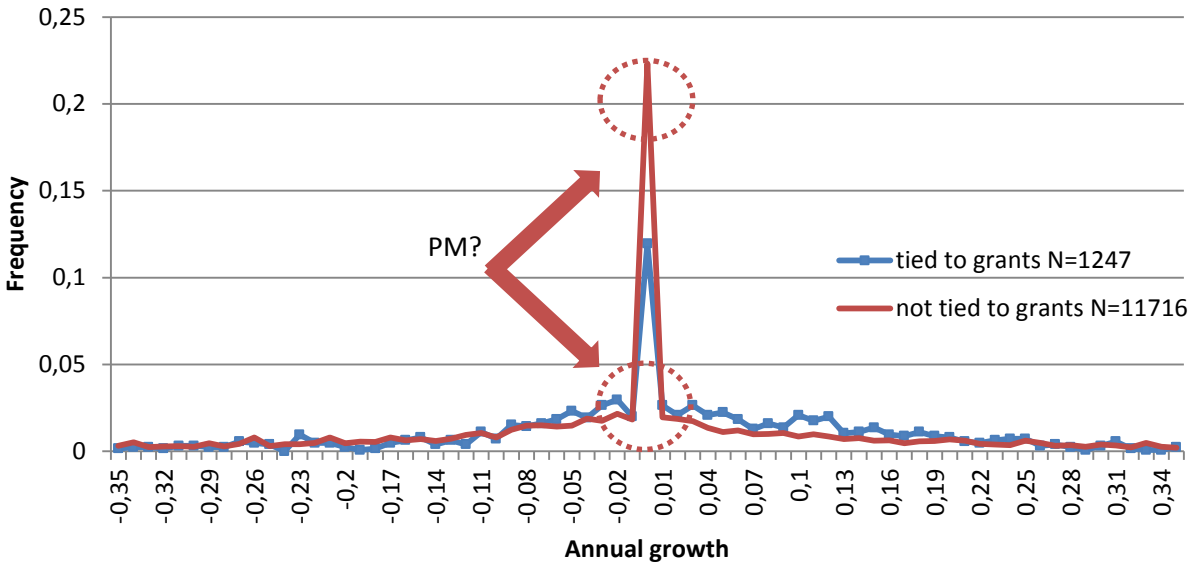
Figure 24 Effect of grants. Traces of prudent (PM) manipulation among not grant-related indicators. Federal ministries. Disaggregated data for 2007-2011.



Self-reported indicators

The effect of prudent manipulation is most visible among self-reported indicators (Figure 25). These indicators show almost no variation from year to year. Growth indices converge to zero for both groups of indicators – tied and not tied to grants. This may indicate that these indicators were most affected by labour-saving. It is possible that instead of collecting actual data civil servants were massaging previous year figures to arrive at the next year values. Only pooled data for five years are reported because there was no visible variation between time periods.

Figure 25. Effect of grants. Traces of prudent manipulation (PM) among self-reported indicators. Pooled data for five years. Growth rates converge to zero in an unnatural way.



The first hypothesis is confirmed. Self reported indicators demonstrate less variation over time. This may be taken as a sign of prudent data manipulation (Figure 15).

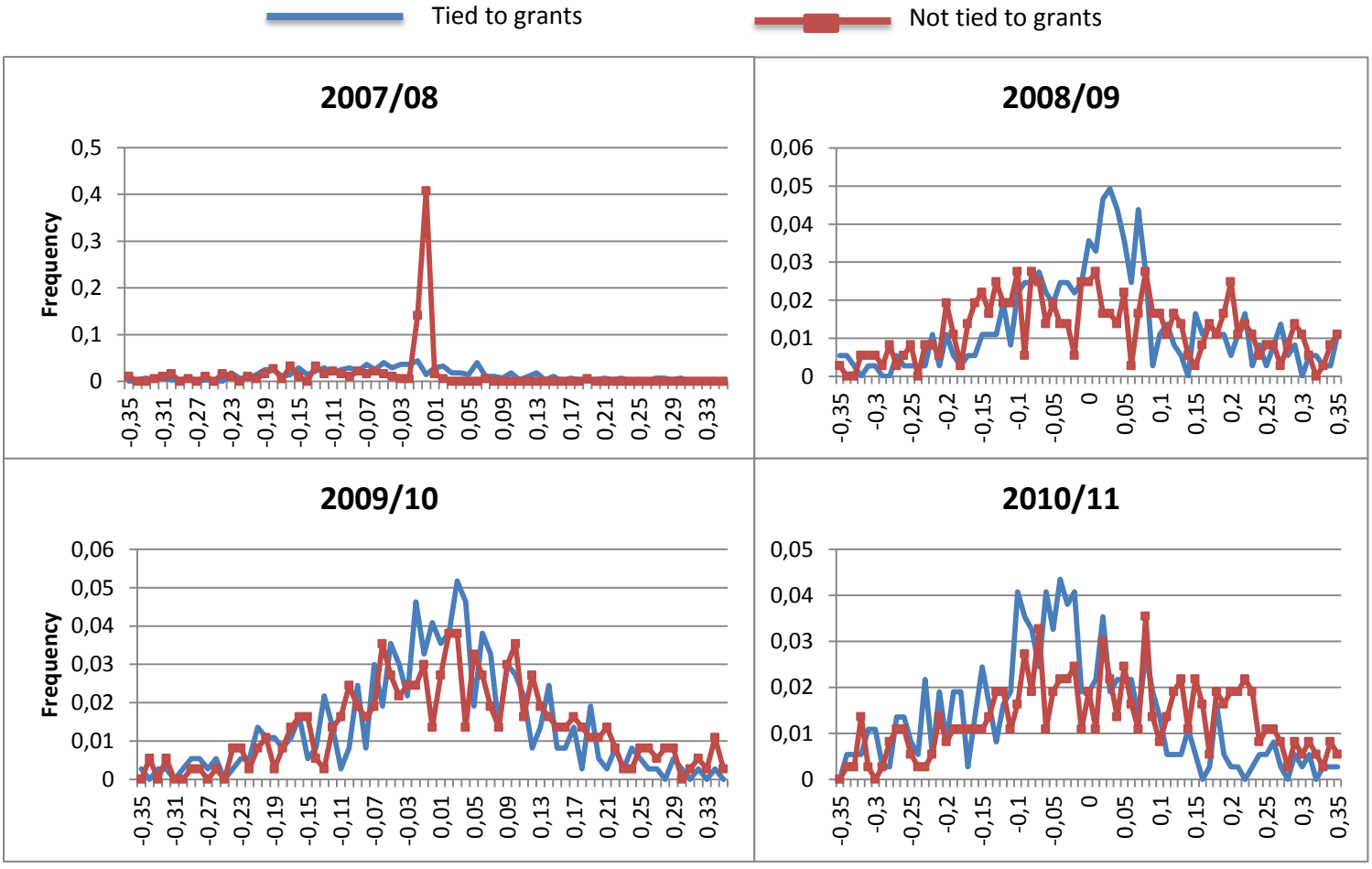
The second hypothesis should be rejected, self-reported indicators show no sign of upward manipulation (Figure 25).

Data reported by Rosstat supports the third hypothesis: grant-related indicators demonstrate more positive performance over time. This may be interpreted as a sign of more reckless manipulation (Figure 21, Figure 22).

The fourth hypothesis is confirmed, indicators not tied to grants appear to be affected by prudent data manipulation (Figure 23, Figure 24).

In addition, an interesting picture may be observed in the case of the Federal Security Service (FSO) (Figure 26 below). In the first year there is almost no variation, the pattern is similar to that of self-reported indicators. In subsequent years the pattern is radically different – indicators demonstrate normal variation between time periods (both tied and not tied to grants). This may be explained by recalling a remark by one of the respondents from a ministry of economic development. He commented on the fact that citizen satisfaction indicators were collected by FSO: "*When we (the ministry of economic development) used to collect citizen satisfaction figures in the Department of Health, we used to get 90% satisfaction all the time. Once FSO took over, they started showing 35%".*

Figure 26. FSO. Citizen satisfaction indicators. When FSO took over the pattern of variation radically changed, suggesting that prior to 2008 data were self-reported and manipulated prudently.



Revision effect

In addition to the effects that were deliberated sought to test the hypotheses, another interesting effect emerged that I named the "revision effect". The effect was exhibited by the group of eight indicators that "changed hands" in 2009. The data for them were collected by regional governments before this year and by federal ministries after this year. This "shock" led to a significant deviation in dynamics in 2009. Indicators were corrected downwards as if the newly assigned agency carried out a revision and decreased the values of the indicators. This behaviour could be interpreted as an example of detected manipulation and an effort to remedy its effects. Curiously, the year after the revision, the indicators began exhibiting the same dynamics as they did before the revision. The graphs below show the effect of revision 2009.

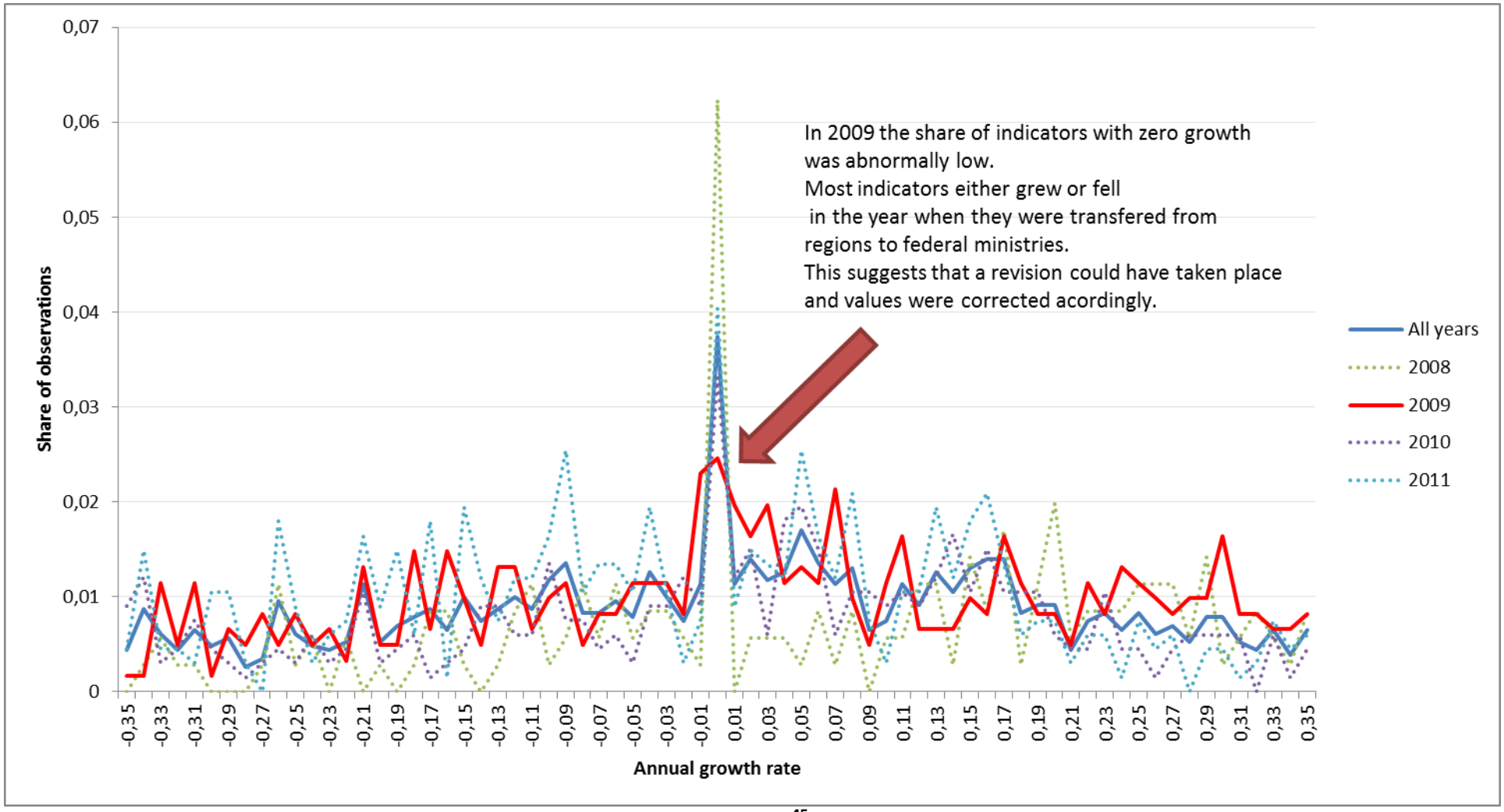


Figure 27. Dynamics of eight indicators that "changed hands" in 2009[45]

[45]Annual growth indices calculated in the same way as in Fig.4. Number of observations within -1…+1 growth rate range: Total - 2295; in 2008 - 353; in 2009 - 609; in 2010 - 664; in 2011 - 669

Alternative explanations for "reckless" manipulation

As for "reckless" manipulation that could be most evidently observed in data reported by the Federal Bureau of Statistics there is a number of possible explanations. First, it is possible that the observed behaviour of indicators tied to grants is the result of reverse causality. Theoretically, grants could be tied to indicators covering areas where improvement of performance was more easily achieved. Qualitative evidence from interviews provides no indication that this was the case. Additionally, cost indicators and payroll indicators were also reported by Rosstat. Annual pay increase could create a visible effect despite correction for inflation. Second, it is possible to explain observed dynamics by genuine improvement of performance induced by the introduction of league tables. However, qualitative evidence speaks against such interpretation. All respondents were sceptical and even cynical discussing the role of performance indicators. For example one regional minister of economic development summarized his experience of dealing with the system of performance measurement as follows: "*I have a feeling that these are just documents for documents' sake… they are not linked to budgeting… their managerial value is absolute zero"*. On the other hand, the major decline in reported performance in 2009/10 may be associated with global economic crisis. If this is the case then Rosstat's data should be credited with reflecting genuine changes in performance.

Results – survey of municipalities

Triangulation with the survey of municipalities

A survey was conducted to estimate the perceived trustworthiness of municipal performance data. Civil servants were asked the following question: "Do you agree with the following statement – "Other municipalities of my region provide reliable and accurate data on values of this indicator". The following choices were given: Strongly agree; Agree; Disagree; Strongly disagree; Do not know.

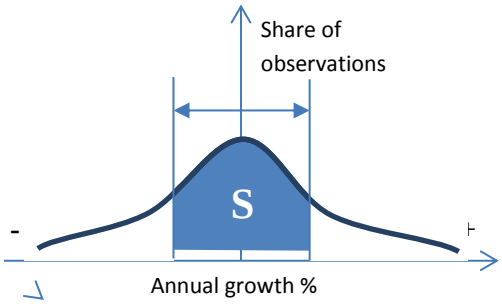
177 answers were collected. 35 indicators were evaluated. Two models were estimated: 1) aggregated answers model and 2) disaggregated answers model. In the former an index of

trustworthiness was calculated by giving 2 points for each "strongly agree" answer, 1 point for "agree", 0 for "do not know". -1 for "disagree" and -2 for "strongly disagree"; in the latter model, no such index was used and disaggregated results were input in the regression for estimation.

The measure of clustering

A "measure of excessive clustering" was constructed to identify indicators that were most likely affected by "prudent" manipulation[46]. Indicators were ranked on this measure and top 20% were marked as "suspicious". See figure below.

Figure 28 Aggregate share of observations in an interval.



Each of 33 indicators was thus assigned a measure of clustering.

A simple linear regression (OLS) was used to estimate the link between the measure of excessive clustering and the measure of trust in performance data. As expected, the coefficient is negative, indicating that higher measure of clustering is, indeed, associated with lower level of trust in performance data.

Aggregated answers model

The model provides statistically weak but visually persuasive results that are easy to interpret. The correlation between trust and clustering is negative as predicted by the "prudent

[46] The measure was calculated as the average of shares of observations (values of growth indices) that fell into intervals around zero (from -0,05 to 0,05, from -0,04 to 0,04, from -0,03 to 0,03; from -0,02 to 0,02, from -0,01 to 0,01; and 0) (Figure 28).

bureaucrat" hypothesis. The measure of trust is constructed by giving the following values to responses: "Strongly disagree" = -2; "Disagree" = -1; "Agree" = +1; "Strongly agree" = +2.

Variable Processing Summary		
	Variables	
	Dependent	Independent
	R5_trust_in_others_data	I_clustering
Number of Positive Values	**33**	**33**

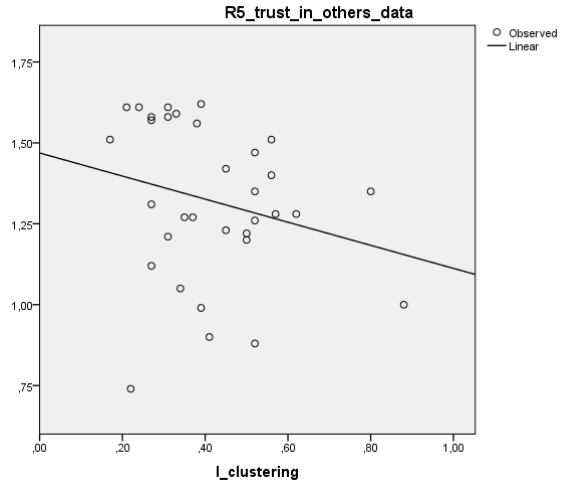
Model Summary and Parameter Estimates

Dependent Variable: R5_trust_in_others_data

	R Square	F	**Sig.**	Constant	**b1 (Coefficient)**
Linear	**,057**	1,884	**,180**	1,469	**-,357**

The independent variable is I_clustering.

Figure 29. Linear fit - trust_in_others_data, i_clustering



Ordinal regression model

Ordinal regression was used to estimate the link between the measure of clustering and trust in performance data. This was done because a 4-point Likert scale was used to capture the degree of trust in data for each of 33 indicators (so the variable is ordinal rather than nominal). Answers were coded from 0 for "Strongly disagree" to 3 for "Strongly agree".

Case Processing Summary

		N	Marginal Percentage
Answers_recoded	0,00	19	0,7%
	1,00	100	3,7%
	2,00	1018	37,2%
	3,00	1600	58,5%
Valid		2737	100,0%
Missing		0	
Total		2737[47]	

Parameter Estimates

		Estimate	Std. Error	Sig.
Threshold	[Answers = 0,00]	-5,176	,272	,000
	[Answers = 1,00]	-3,303	,173	,000
	[Answers = 2,00]	-,541	,149	,000
Location	**I_clustering**	**-,852**	**,246**	**,001**
	I_rapid_growth	1,036	,519	,046

Link function: Logit.

The model supports our hypothesis : greater degree of clustering in associated with lower trust. The results of the survey may be used for triangulation to corroborate the "prudent bureaucrat" hypothesis. Statistical significance is at 0,1% level. Indicators that were most likely "prudently" manipulated appear to be less trusted by civil servants. It may be that they feel the weakness of the data from their first-hand experience of working with these indicators themselves.

[47] The total number of observations. This number is arrived at by multiplying the number of rated indicators (33) by the overall number of responses. The number of responses varied across indicators providing the average of 96 responses per indicator giving 3168 responses including "Do not know" or 2737 excluding "Do not know".

The other hypothesis ("reckless" manipulation) appears not to be corroborated. Indeed, it appears that the correlation between the measure of trust and the "index of rapid growth" is positive. Better performing indicators are more trusted.

Limitations of statistical analysis

It should be emphasized that the statistical method used in this chapter to identify suspicious performance data is likely to be of limited value in identifying individual cases of fraud. It is useful in identifying overall bias in incentives, but it cannot predict that any given observation is fraudulent. The usefulness of the method is limited to systems design.

Similarly to the use of digital analysis in auditing (such as the Benford's law test (Durtschi et al. 2004)) statistical analysis should be used as a first step that enables to broadly identify suspicious data. Closer examination of individual indicators and regions should follow to identify fraudulent practices.

However, statistical analysis is useful in analysing the structure of incentives created by performance reporting procedures. In this case, there may be no intention to uncover individual cases of fraud, but rather to estimate overall effect of existing incentives on the process of performance reporting.

An attempt to separate cheating and "natural" performance

One of the main limitations of quantitative analysis is that it is impossible to separate effects of data manipulation from effects of increased focus on the indicator and other types of unintended consequences. Quantitative findings presented in this chapter demonstrate that self-reported indicators were concentrated in the vicinity of zero to a much greater extent than indicators reported by other agencies. How much of this effect is due to manipulation, gaming or the natural propensity of the measured phenomena to stagnate despite any government interventions. The problem here is that all that we know is the distribution of data *after* supposed manipulation. Quantitative analysis of existing data does not let us separate the "natural"

variation of underlying phenomena from effects that occur during and after measurement. We made an attempt to estimate the "natural" variation by asking civil servants to estimate the likely behaviour of performance indicators in the absence of government intervention.

As part of the aforementioned survey an attempt was made to assess the "natural" variation of performance indicators. Civil servants were asked the following question: "What dynamics would the indicator show in the absence of any government intervention?". The underlying logic was that by asking this question it would be possible to roughly assess the "natural" variation of underlying phenomena. This is essential to separating the effect of data manipulation from the natural propensity of the indicator to fluctuate in a particular way.

The results show that civil servants do not perceive any considerable difference between indicators reported by different agencies. So the *ex ante* "natural" variation is not significantly different (see table below) 27% for Rosstat vs 21% for regional government's indicators. *Ex post* variation, is, however, is significantly different (8% for Rosstat and 23% for self-reported indicators (the share of observations with zero growth) (Figure 15). This may be interpreted as a hint at the scale of manipulation. Inherently the areas of public service measured by self-reported and other indicators are not perceived as significantly different, but after measurement they demonstrate significantly different dynamics. More details are given in Table 16Table 16.

Table 15. The tendency of indicators to fluctuate in the vicinity of zero growth in the absence of government intervention (percentage of respondents)

Reporting agency	"the indicator would moderately fluctuate in the vicinity of zero growth"*
Fed. Stat. Bureau	27%
Fed. Ministries	22%
Municipalities	21%
Regions	21%

* average number of responses - 96

Table 16. Assessment of the dynamics of indicators in the absence of government intervention (Q2)

	In the absence of government intervention this indicator would…*							
Reporting agency	grow slowly	grow quickly	drop slowly	drop quickly	**fluctuate moderately around zero**	fluctuate significantly around zero	Do not know	No. of indicators
Fed. Stat. Bureau	23%	4%	24%	8%	**27%**	5%	10%	14
Fed. Ministries	24%	7%	21%	9%	**22%**	3%	15%	16
Municipalities	25%	10%	21%	4%	**21%**	2%	17%	13
Regions	14%	5%	27%	11%	**21%**	2%	22%	13

* average number of responses – 96

Measuring other "natural" characteristics of performance indicators

Table 17 Q1. In your opinion, how significant is the influence of municipal governments over a given indicator? (0 to 100%)

Answer	Rosstat	Federal ministry	Municipal government	Regional government
0% (the dynamics of this indicator is entirely determined by other public sector agencies)	**15%**	8%	6%	7%
25%	20%	10%	9%	10%
50%	20%	17%	16%	14%
75%	13%	18%	17%	16%
100% (the dynamics of this indicator is entirely determined by the municipal/regional government)	**24%**	43%	42%	45%
Do not know	8%	5%	9%	8%

In line with expectations indicators collected and reported by the federal statistics (Rosstat) bureau have a lower score. They are perceived to be less under the control of municipal/regional civil servants. Strangely, indicators reported by federal line ministries have similar scores to locally generated indicators. This may mean that there are close ties between federal line ministries and regional administrations than between regional administrations and Rosstat.

Table 18. Q3. In your opinion, if the government had abundant resources and could allocate them to the area measured by this indicator, how easy would it be to achieve growth of 5% in a year?

	Rosstat	Federal ministries	Municipal	Regional	
0	23%	22%	21%	24%	Very easily achievable
1	57%	50%	51%	49%	Achievable
2	9%	9%	9%	7%	Hard to achieve
3	2%	3%	2%	2%	Not achievable
99	9%	15%	16%	18%	Do not know

This question is an attempt to measure the natural "elasticity" of indicators, that is, the ease of demonstrating significant growth by allocating greater resources to the measured area.

Answers show that there is no perceived difference between indicators reported by different municipalities in the degree of "natural elasticity". Breaking up indicators into four groups according to the reporting agency does not produce any meaningful differentiation. This may be taken as a supporting evidence for the hypothesis that there are no *ex ante* differences between indicators in their propensity to grow more or less quickly in response to government interventions.

Table 19. Q4. In your opinion, is your municipality able to collect objective and accurate data on the value of this indicator? Do you agree with the following statement – "Our municipality can collect accurate and reliable data to measure values of this indicator.

	Rosstat	Fed.min	Municipal	Regional	
0	60%	77%	66%	71%	Strongly agree
1	27%	19%	26%	21%	Agree
2	7%	1%	3%	2%	Disagree
3	1%	0%	1%	0%	Strongly disagree
99	4%	2%	4%	6%	Do not know

This question is an attempt to evaluate the perception of the quality of data available to civil servants for measurement of performance.

The results show that indicators reported by federal line ministries are most trustworthy. But the difference between locally generated (regional/municipal) and federally generated data (Rosstat) is not decisive.

Table 20. Q5. In your opinion, how accurate and objective are the data on values of this indicator, reported by other municipalities in your region? Do you agree with the following statement – "Other municipalities of my region provide reliable and accurate data"

	Rosstat	Federal ministries	Municipal	Regional	
0	49%	58%	49%	53%	Strongly agree
1	34%	32%	36%	32%	Agree
2	4%	1%	3%	2%	Disagree
3	1%	0%	0%	0%	Strongly disagree
99	12%	9%	12%	13%	Do not know

This question is an attempt to evaluate general trust in data reported by other municipalities and whether it varies based on the source of data (reporting agency).

Responses show that there is no decisive difference between the level of trust and the source of data. These results should be contrasted against the econometric analysis provided above. There is a strong correlation between "prudent bureaucrat" manipulation pattern (convergence of growth indices to zero), but, as this question shows, there is no *ex ante* difference in trustworthiness based on the source of data. Not all regional and municipal indicators are perceived as less trustworthy, but only those indicators that tend to demonstrate a particular type of behaviour. This may be taken as evidence that supports our method of detecting "prudent" data manipulation.

Overall the results of the survey of municipalities show that indicators reported by different agencies are not inherently different in terms of trustworthiness or "elasticity" (ease of improving). *Ex ante* differences are not sufficient to explain the variation in trustworthiness correlated with "prudent bureaucrat" patterns in data. In other words, excessive convergence to zero is not an inherent characteristic of a set of indicators but is (at least in part) a result of the process of reporting. Indicators are not seen as dramatically different *ex ante*, but show dramatically different dynamics *ex post*. This supports the hypothesis that cheating/gaming plays a large role in forming the observed patterns.

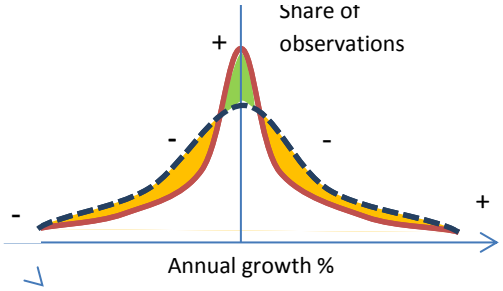
Further research opportunities

For example, areas below the curves may be compared. Table 21 reports the share of observations that fall into progressively narrowing intervals around zero (see Figure 30 for illustration).

Table 21. Aggregate share of observations in intervals around zero.

Interval	Rosstat	Self-reported	Federal Ministries	Federal Security Service
-0,05 - 0,05	0,36	0,40	0,38	0,32
-0,04 - 0,04	0,31	0,37	0,33	0,27
-0,03 - 0,03	0,26	0,34	0,28	0,22
-0,02 - 0,02	0,21	0,30	0,22	0,17
- 0,01 - 0,01	0,15	0,26	0,15	0,11
0 (kurtosis)	0,08	0,21	0,07	0,05

Figure 30. Measure of excessive convergence towards zero - illustration.



By subtracting the control group's values from the values of the evaluated group we can arrive at a measure of difference between the two groups. This measure may be interpreted as the measure of excessive convergence of the group of self-reported indicators. By doing empirical research (cross-country or cross-industry) it may be established what criteria should be used to distinguish between the "normal" or natural level of clustering and the level that indicates deliberate data manipulation.

In our example the following values were obtained (Table 22). An illustration is provided by Figure 30. Green area indicates positive values, orange areas indicate negative values.

Table 22. The difference between evaluated groups and the control group in aggregate share of observations in near-zero intervals.

	$S_i - S_{Rosstat}$		
Interval	**Self-reported**	**Federal Ministries**	**Federal Security Service**
-0,05 - 0,05	0,041	0,018	-0,038
-0,04 - 0,04	0,060	0,020	-0,042
-0,03 - 0,03	0,079	0,022	-0,037
-0,02 - 0,02	0,096	0,015	-0,038
- 0,01 - 0,01	0,108	0,000	-0,040
0 (kurtosis)	0,127	-0,010	-0,033
Summ	***0,483***	***0,066***	***-0,239***

This measure provides a way of quantitatively assessing the scale of "prudent bureaucrat" manipulative behaviour. "Reckless bureaucrat" behaviour may be similarly assessed by measuring the right-hand tail of the distribution (left-hand tail if improvement is represented by negative growth).

In this case the value of 0,483 (or **48,3%**) is associated with a situation where significant incentives for manipulation are theoretically predicted. It is arrived at by aggregating across 75 self-reported indicators from different spheres of government activities (Health, Education, etc.).

Further research is needed to establish criteria for specific spheres, countries and measurement systems and recommend more appropriate measures. Development of such methods may provide a useful way of evaluating performance measurement systems for susceptibility to adverse practices. The process of refining performance measurement systems (or any system in general) proceeds by trial and error. In order to tell success from failure it is necessary to have ways of measuring the scale of the problem that is being treated. Explicit quantitative measures of the scale of manipulative behaviour would allow for evaluation of the system before and after the implementation of any remedial initiatives. Thus, the necessary conditions for cost-effectiveness analysis may be created.

The Russian experience demonstrates that the government have (possibly) recognized the problem with self-reported indicators and abolished them completely in the most recent round of changes of the system. The system, thus, was swung from one extreme to the other and significant amount of locally generated data was abandoned. I believe that a more gradual approach could be taken by reforming the incentives structure and evaluating the effectiveness of step-by-step changes. The methods outlined in this section could facilitate a more informed, rational and gradual approach to reform.

Conclusion

Quantitative findings suggest that the two strategies of data manipulation, identified based on respondents' accounts, may be responsible for a noticeable bias in trends of performance indicators. The type of a principal-agent model most accurately describing the relation between regional governments and supervising federal body seems to be associated with different patterns in performance data. Self-reported indicators tend to be more stagnant over time but this is in line with the "prudent bureaucrat" hypothesis. Evidence for "reckless manipulation" is inconclusive. Somewhat surprisingly it turned out that indicators report by Rosstat seem to be affected by "prudent bureaucrat" manipulation – noticeable drop in small negative values suggests that data were "massaged" upwards.

These interpretations of observed dynamics of indicators appear more plausible against the background of overall scepticism towards the existing system of performance measurement. Qualitative findings presented in the previous chapter showed that respondents did not see instances of data manipulation as abnormal exceptions to the rule, but rather as the rule itself. Their attitude cannot even be characterized as ambivalent, because there seemed little if any condemnation of the practice of "painting" (or "drawing") performance figures. It is as if lower rank civil servants were united in their disrespect for the system and the top level officials did not regard the system important enough to spend their time on improving its accuracy.

The situation was not helped by the ever-growing number of indicators that were collected and reported within the framework of the system. The number of indicators increased from just over 40 when it was initially formulated by the president in 2007 to 325 in 2012 when the system was discontinued. Such an avalanche of performance indicators meant that the system either had to become too expensive in terms of time and resources, or ways of minimizing effort had to be found. One of such ways was the "drawing" of performance figures that allowed civil servants to save the hassle of actual collection and processing of performance data.

These considerations make it possible to give an explanation of the observed anomalies in dynamics of performance indicators outlined in this chapter. Since quantitative results were obtained on the sample covering the entire population (83 regions, 325 indicators, 5 years), it is possible to generalize the findings beyond the boundaries of the initial selection of three regions. The following chapter attempts to provide theoretical generalization and a set of possible causal mechanisms that could explain the emergence of observed phenomena.

CHAPTER 7. DISCUSSION

The findings presented above contribute to different streams of literature. First, there is the discussion of types and methods of deliberate data manipulation in both public and private sector. Second there is a wider discussion of unintended consequences of performance measurement systems in general. Third, there is a more general discussion of incentives that emerge in bureaucratic hierarchies and how these change or are changed by modern approaches to public sector management. Then, fourth, there is the discussion in development literature on home-grown versus imported policies in public administration. Finally, there is distinct literature on Russian regional politics and public administration.

Smith (1995) identified 8 types of unintended consequences that result from introducing formal performance measurement. Some of these are related to "real" managerial effects on behaviour such as concentrating on measured aspects of a service at the expense of unmeasured ones. Other unintended consequences are related solely to data generation process: performance may be misrepresented in various ways. However, these assertions tend to be supported exclusively by anecdotal evidence. Quantitative methods are rarely employed to demonstrate the scope and pervasiveness of data manipulation. Partly this is due to the difficulty of separating effects of deliberate data manipulation from effects of other factors. I suggest that the system of performance measurement that operated in Russian regional governments between 2007 and 2011 allows for quantitative analysis because of some of its unique features: 1) public availability of the data; 2) existence of clearly definable groups of reporting agencies with varying incentives to manipulate the data; 3) no provisions for routine formal audit; 4) no provisions for formal sanctions for data manipulation; 5) the "hermetic" nature of modern Russian bureaucracy that insulates it from public scrutiny and largely from political interference at local and regional levels.

Limitation of the research

Qualitative study

The study is limited in its generalizability. The number of interviewees (25) and the number of regions (three) is limiting. The number of respondents was sufficient to reach the saturation point of in-depth interviews, but a larger number of respondents from a more diverse selection of regions could provide further findings (for example, other strategies of data manipulation could be discovered). Similarly, the fact that respondents were selected mainly from middle level civil servants could generate a bias. A more diverse selection of respondents could generate additional insights.

Quantitative study

The methods used during the quantitative stage also have their limitations. Observed irregularities in distribution of performance data could be the result of some inherent difference between public services. The main difficulty here is separation of the results of data manipulation and honest efforts at improving performance. Although an attempt was made during the survey of municipalities to tackle this limitation, it remains questionable whether the effects of cheating can be separated from natural trends in the data.

The main weakness of the survey of municipalities was that respondents were given a set of hypothetical situations to choose from. Since respondents were municipal workers it may be problematic to generalize their answers to the regional context. Still it was the only practically available opportunity of assessing trustworthiness of performance data.

A note on the unit of the analysis

Ludwig von Mises in his *magnum opus Human Action* advocated the use of the term "praxeology" to speak about the science of human action. As far as I understand, his use of the term may be understood to mean something like (to use a more widely used term) phenomenology of human action. Praxeological knowledge is derived not from observing

empirical data but from *a priory* reasoning about human action *per se*, abstracted from all empirical contingent content. If my understanding of the term is correct, then we can apply praxeological reasoning not only to economic phenomena as Mises did, but also to other spheres of human action. Particularly, we can try to study systems of rewards and sanctions in organisations using praxeological reasoning. It is beyond the scope of this study to develop this approach, but it is useful to consider what seems to me the basic elements of praxeological understanding.

Praxeologically (or phenomenologically) what is called "performance indicators" in the UK is not the same entity (practice) that is called "performance indicators" in Russia. The system of rewards and sanctions that creates the praxeological category "performance indicator" is dissimilar in the two contexts. This difference stems from deep dissimilarities in all material and ideal existence of civil servants in the two contexts. Wage rates relative to hours worked are different, technological level is different, political system, economic situation, culture, language, mentality, etc. are all different. To say then that two managerial tools are the same in the two contexts is to say that they generate similar praxeological (or behavioural) responses from civil servants in the two contexts. But this appears as impossibility, because merely giving a tool the same name as that of its foreign prototype is not sufficient for shaping the network of practices that gets attached to the tool. This may be illustrated by comparing the importance of informal rules in the two countries.

In the public sector in Russia the practice of civil servants is in a very large portion regulated by informal tacit rules of interaction. This is not to say that there are no informal tacit rules in the British public sector, but I think that it would be hard to dispute the assertion that the relative importance of formal rules is higher in Britain than it is in Russia. Moreover, informal rules in Russian and in Britain are different. Therefore, any practice that is implemented in Russia necessarily has to adapt to the network of tacit rules and norms of the public sector and the culture in general. Russian culture has been described as "context rich" culture. In addition to

this Russia has a different political regime (it has been described as “Limit Access” regime in contrast to “Open access” regimes in developed western democracies (North et al. 2009). Moreover, politically Russia is a peculiar sort of hyper-centralized federation.

These differences mean that, if one attempts to implement two formally identical practices (say, performance pay) in the two contexts, they would get embedded in two different networks of practices and would be shaped by existing field (or plane) of practice. Despite having two identical names and formal descriptions these two reforms would, from the very start, exert different influence on individuals because they would operate through and by the means of different praxeological categories. Attitudes towards them would be different on a fundamental level of practice in general. Human action would be changed differently by these two seemingly identical practices, because in reality they are not identical. It may be argued that it is merely a difference in degree, not in kind. But I would argue that even if it is a difference in degree (and it is disputed whether subjective feelings may have differences in degree, then the accumulated effect of differences in degree in the rich network of practices would mean that the new entity itself would generate different “praxeological gravity” and shape human action in a different way. A metaphor may be used to illustrate this point. In physics, bodies that have large mass change the curvature of space-time (they reshape the fabric of space-time continuum by their gravitational field in such a way that other bodies change their trajectories of movement). Similarly in human practice one practice may have large gravity and may shape the flow of other practices. These practices include managerial, technical, political and other types of interactions between government bodies, between different tiers of government, between individual civil servants and between public sector organizations and wider groups of shareholders.

When a new practice emerges (say, a routine of publishing performance indicators) it is akin to appearance of a new massive body in the space-time continuum. This new body gets enveloped in the existing network of forces and simultaneously changes them and is affected by

them. The resultant characteristic of the new state of rest depends both on the characteristics of the new body and on the pre-existing characteristics of the network of forces.

It is dubious whether it is possible to strictly speak of "similar functions" performed by two systems in different contexts. Empirically the appearance of functions may be similar, but they praxeological content may be dissimilar and vice versa. So one needs to decide what one means when talking about "similar functions". Similar praxeological functions in Russia might be performed by some different technocratic instruments (it has been argued, for example, that the function of assessment of regional administrations is performed by electoral results of the ruling party. Percentage of votes received by United Russia was used to evaluate the performance of governors[48]. This "integral indicator" foreshadowed all formal evaluations and performance management practices).

So it is hardly possible, strictly speaking, to compare system of Public Service Agreements in Britain to PM systems in Russia, because they produce different praxeological effects, exist in different contexts and get effected by different practices.

Thus, the selection of the unit of analysis is inherently problematic. Merely comparing two systems with two similar names ("performance indicators") is not satisfactory because under similar names there exist two different networks of practices. One has to say that the unit of analysis is the locus of practices that emerged when an attempt was made to introduce a system with similar basic characteristics. This necessarily means that the borders of such unit need to be blurry and somewhat ambiguous. Because of the interference of other practices and imbeddedness in local context different elements may be included in the analysis. For example, political systems is rarely discussed in managerial studies of performance management regimes in western democracies. It would, however, be short-sighted to avoid discussing political implications of the system of performance management in Russia where the structure of the "vertical line of power" exerts its gravity on any managerial tool. In Russia governors were

[48] Yakovlev A.A. 21.02.12 "Ne tot signal podali gubernatoram" (The wrong signal to governors) http://slon.ru/russia/ne_tot_signal_podali_gubernatoram-749143.xhtml

appointed between 2004 and 2012, and were elected before this period. It would be unwise to ignore this fact in discussing the evolution of national performance management regime.

Theoretical generalization from the Russian case

As discussed above, Russian bureaucracy is less burdened with layers of checks and balances than public bureaucracies of the Western world. It is less subordinated to the control of civil society. When formal transparency is introduced in such a system it becomes accessible for research without immediately losing its natural habits and character. I think it is, therefore, appropriate to apply to the Russian case some of the theories of bureaucracy that were developed in the second half of the XX century and have by now become classic. More recent developments (such as new public governance, network governance etc.) seem less appropriate because they emerged to tackle issues of public administration that appeared in developed western democracies and have yet to take root in Russia. At the same time, findings of this project may be of interest for practitioners in developing countries who may find them applicable and theoretically appropriate for conditions of their countries.

The following theories are used for theoretical generalisation: theories of bureaucracy by Gordon Tullock and Antony Downs, the model of administrative man by Herbert Simon and Michel Foucault's notion of normalisation.

The natural selection in the "vertical line of power" as a filter against honesty

The "vertical line of power" is a term coined by Vladimir Putin to refer to the highly centralised structure of the executive branch of power. The creation of this vertical line of power may be regarded as one of the achievements of the Putin era. This centralised structure may be studied using classical theories of bureaucracy by such authors as Gordon Tullock and Antony Downs. In his influential work on the theory of bureaucracy published in 1965, Gordon Tullock (2005) used an analogy of a gaseous diffusion plant to illustrate the process of career selection in any hierarchy based on merit. Individuals entering the system are continuously tested and either

rise, fall or remain in place as a result of the test. Only those who systematically make decisions in the interest of their career are likely to rise to the top:

> *Any political hierarchy in which personnel are selected for promotion by the system we have designated "merit" will function in much the same way. People entering the system are either a random selection or the result of a preliminary selection process. Once they are in the system, they are confronted with a number of situations in which they may either rise, remain in the same position, or fall. These "test" situations do not necessarily refer to formal promotions in the bureaucratic hierarchy. There are usually numerous smaller steps which prepare the way for formal promotion or demotion. The obtaining of a good assignment, earning the confidence of your superiors, getting a "good name around the office," all may be equated to the porous barriers of the gaseous diffusion plant. (Tullock and Rowley 2005: 21)*

…

> *Any individual in any system will continually be confronted with choices between courses of action which will have at least some favourable effect on his chances and others which are less desirable from that point of view, but which have other advantages. Only the person who usually chooses in terms of his "career" will be likely to rise to the top.(Tullock and Rowley 2005: 22)*

Since people can adjust their behaviour to the rules of the game and since individual objectives often differ from organizational objectives, such a system inevitably selects against honest individuals and promotes those who always pursue the course of action that is favourable for their personal advancement even at the detriment of organizational objectives:

> *It is impossible to design a system that will select against the man of relatively low morals. This is because the intelligent but unscrupulous man will always assume the morally proper course of action if, in fact, this should be the one that is the most likely to be successful (Tullock and Rowley 2005: 26).*

An "idealistic" individual who attempts to pursue "the general good" will find himself in a disadvantage when confronted with a choice between two actions one of which is likely to bring him a career promotion but is bad for the organization as a whole, and the other one which is in true interests of the organization but is likely to hurt the individual. Thus, "any organizational structure in which selection on a merit basis is employed is likely, at least to some extent, to select against morality" (Tullock and Rowley 2005: 26).

A "realistic" individual, who habitually chooses in his individual interests, may not necessarily be aware of the damage he is inflicting:

> *It is always difficult to distinguish between "what is good for me" and "what is good." The general good is never readily discernible. The "politician," the bureaucrat, who makes no especial effort to keep these two categories distinct can quite genuinely believe that a course of action which may appear cold-blooded and dishonest to the outsider falls legitimately within his range of duty. (Tullock and Rowley 2005: 27)*

In an ideal organization one never has to choose between "what is good for me" and "what is good". It is unlikely, however, that his situation is common:

> *In the ideally efficient organization, then, the man dominated by ambition would find himself taking the same courses of action as an idealist simply because such procedure would be the most effective for him in achieving the personal goals that he seeks. At the other extreme, an organization may be so badly designed that an idealist may find it necessary to take an almost completely opportunistic position because only in this manner can his ideals be served. The idealist, in such cases, may find that only by taking the course of action that will advance his own career can he remain in the organization and advance to a position where he can hope to influence events. This is administrative organization at its worst (Tullock and Rowley 2005: 24)*

…

> *The general "moral level" of those bureaucrats who have reached the top layers in such a structure will tend to be relatively low". (Tullock and Rowley, 2005:25)*

The public sector is known to be the place where one inevitably faces such dilemmas. Thus, naturally moral people may prefer to avoid public sector employment deliberately. A mechanism of negative self-selection is then set in motion. As a result, "few people expect career civil servants to act contrary to their own interests" (Tullock and Rowley 2005, 27).

In Russia the situation may be exacerbated by the dominance of informal rules, ambiguity in criteria for career promotion (Yakovlev 2010, 24), high turn-over and low organisational loyalty (Gimpelson, Magun, and Brym 2009).

A recent survey of civil servants[49] found that over 73% of civil servants think that the most relevant criteria of career promotion is the disposition of their immediate supervisor[50] and a large

[49] http://cinst.hse.ru/contract (publications forthcomming)

proportion of civil servants find criteria of career growth unclear and unpredictable (for a detailed analysis of the system of promotions and entry selection in the Russian public sector see (Gimpelson et al. 2009)) .

It is, therefore, reasonable to expect that any ambitions civil servant, when confronted with the choice of whether to manipulate the data would choose to please his immediate supervisor by showing figures that show him and his department in a more favourable light. Since the criteria of career promotion are unclear and all depends on the discretion of supervisors, it is rational to try to portray a good image "just in case":

> *The paradox is that if one tries to introduce formal indicators without clearly defining their role and status, a system of informal incentives spontaneously emerges (including the desire to get noticed, etc.). The resulting effect of such indicators on the system is hard to predict, account for and correct. (*Jakobson 2006, 19*)*

The system of incentives tied to performance reporting was unclear. Apart from formal grants to top performing regions, which respondents considered a minor decorative feature, the system lacked defined rewards and sanctions. However, it was implemented in the context where informal rules were paramount and, thus, it became affected by traditional bureaucratic careerism.

When facing with a choice of whether to report a discovery of unrealistically inflated figures (perhaps, done by his predecessor or by another agency), it would take exceptionally courage and moral fibre to blow the whistle and call for an investigation. One who decides to uncover such manipulations may expect his career prospects to be worsened[51].

---

[50] One of the questions read: "Who, do you think, is in the position to most accurately judge the results and quality of your work?". Responses: 73% immediate supervisor; 25% the head of department. Another question: "How clear and predictable are the criteria of career growth in the civil service?" Responses: clear and predictable – 19%, all depends on the immediate supervisor – 24%, all depends on the head of department – 21%, criteria are unclear and unpredictable – 19%, do not know – 17%.

[51] An old Turkish joke illustrates the point well: "Once a long-serving head of province suddenly died and a young inexperienced official was appointed to replace him. Soon an order came from Istanbul to provide an updated report on the number of trees growing in the jurisdiction. The official duly sent his staff to count the number of trees and reported the number to his superiors. After some time an urgent request came back from Istanbul, it read: "Immediately explain why half of all the trees in your province disappeared?!"

If this situation is present, the system may become self-destructing in the long run:

> *…the barriers act so as to select by criteria that are not only irrelevant from the standpoint of the designers, but which will, in the future, result in even poorer performance and selection.(Tullock and Rowley 2005: 24)*

If civil servants are in a position to manipulate data unchecked, the whole edifice of performance management built on these data becomes shaky as the data increasingly detaches from reality.

It is not certain how far the data diverged from reality as a result of deliberate manipulation, but it is informative to consider the fact that in 2012 the system of performance measurement in Russia was radically changed. Out of 325 indicators only 60 remained and out of these 60 not a single indicator was to be self-reported by the regional administrations.

It appears that the presidential administration learned the lesson and discontinued the bulk of indicators that became largely meaningless. One factor of this could be the two data manipulation strategies outlined above.

In addition to Tullock's theory of negative selection another theoretical perspective may be employed to explain the use of "prudent" manipulation strategy. This strategy, presumably was associated with the desire of civil servant to 1) keep a low profile; and 2) minimize efforts and labour spent on performance reporting. Both desires are accounted for by the model of "administrative man" proposed by Herbert Simon and the notion of "normalisation" elaborated by Michel Foucault.

Linking data manipulation strategies to bureaucratic personality types

The results may be theoretically generalised using the theory of bureaucratic personality types developed by Antony Downs (1967:85). Downs identified two types of motives making up bureaucrats' "utility functions": self-interest and altruistic (broader) motives. Self-interested motives include *power, money income, prestige, convenience and security*. Broader motives

include *personal loyalty, "mission-commitment", pride in proficient performance of work and desire to serve "the public interest"*.

Using this typology of motives Downs defined five bureaucratic personality types: two "purely self-interested" and three "mixed-motive" types (Downs 1967, 88). The two self-interested types – *climbers* and *conservers* – fit well with the two observed strategies of data manipulation (Table 23). Presumably, data manipulation strategies used by bureaucrats driven by broader motives would be more subtle, context-specific and difficult to identify.

Table 23. Personality types of bureaucrats adapted from (Downs, 1967, p.88)

Personality type	**Dominating motives**	**Comment**
Purely self-interested officials		
Climbers	power, income, and prestige	
Conservers	convenience and security	In contrast to climbers, conservers seek merely to retain the amount of power, income, and prestige they already have, rather than to maximize them.

Bureaucrats belonging to different personality types, according to Downs, have different value systems and rank their motives differently. This difference may result in resorting to different courses of action when an opportunity comes to engage in data manipulation.

Conservers would aim at minimising risks and maintaining *status quo*. They would prefer to play safe by resorting to the "prudent bureaucrat" strategy. It is prudent from a *conserver's* point of view to avoid consistent inflation of performance figures. A *conserver* is motivated by the desire for maximum long-term security and minimum inconvenience. Thus, it is preferable for him to be inconspicuous. Any above-the-average as well as below-the-average performance is risky as it may attract superior's attention.

Climbers, on the other hand, would aim at maximizing their career prospects, by showing their agencies' performance in favourable light no matter the risk of sanctions. If faced with the same situation, they are likely to engage in deliberate inflation of their performance figures

following the "reckless" manipulation strategy. This is likely to lead to severe penalties if found, but, if one is successfully in concealing this fraud, one may earn a promotion or obtain other perks.

Downs's theory of bureaucratic personality types may be usefully applied to a wide range of organisational contexts. If there is a link between a bureaucrat's personality type and a strategy of data manipulation he/she prefers then potential manipulations can be productively predicted from the knowledge of prevailing mentality in a given organisational context. And, vice versa, the information about the dominant strategy of data manipulation may be useful in diagnosing the dominant type of mentality in an organisation. It appears that in the Russian case prudent manipulation was more widespread. It was especially pronounced in the set of indicators not linked to grants.

Tullock's theory of negative selection in hierarchies and Downs' typology of bureaucrats are particularly relevant for the Russian case. Indeed, the Russian public sector is facing important ethical challenges:

> *The Russian Federation is currently in transition from an ideologically driven society based on social class to one in which ethics are based on individualistic and civic principles. (Barabashev and Straussman 2007: 380)*

In such transition ethical safeguards are temporally weakened and the problems of ethics and morality are exacerbated:

> *Thus, ethical safeguards, which are supposed to prevent antisocial, professional behaviours and corruption, are either weakened or absent. They have been replaced by short-term self-interest incentives. The new mechanisms of the merit-based system are only beginning to be realized in Russia. Presently, the Russian bureaucracy lacks unity because personal interests dominate the public interest. (Barabashev and Straussman 2007: 380)*

In this context data manipulation in personal self-interest is more likely than in the contexts with strong established values of public sector honesty and integrity. Indeed, this view is shared by Russian politicians:

President Putin described the Russian bureaucracy as a closed and haughty caste system that sees public service as a kind of business rather than as an institution responsible for serving the citizenry.(Barabashev and Straussman 2007: 376)

Linking data manipulation strategies to types of performance measurement regimes

Hood (2007b) offers a typology of performance measurement regimes. He classified performance measurement regimes into *target, ranking* and *intelligence* systems. *Target* systems aim at aggressively driving change by setting minimum standards or aspirational quantifiable goals and establishing sanctions and rewards. *Ranking* systems use league tables to identify winners and losers and apply appropriate measures. *Intelligence* systems aim at providing background information to facilitate learning without added pressure of formal targets and rankings.

From this perspective, the Russian system of regional performance measurement may be characterized as a hybrid system of *intelligence* and *ranking* indicators. Within one framework some indicators were formally used for ranking, others were only reported for intelligence purposes. Grants were allocated on the basis of the league table to top 10 performing regions (league tables were calculated using 63 selected indicators, these indicators were included in the formula for federal grants).

In the previous chapter it was demonstrated that among indicators included in the league table, signs of "reckless" manipulation were observable. In contrast, indicators not included in the grant formula were exhibiting a different trend - the dynamics in accord with the "prudent" manipulation strategy. This may indicate that "prudent" manipulation is a characteristic feature of *intelligence* indicators. They do not provide sufficient incentives for civil servants to engage in ambitious manipulation. These *intelligence* indicators are labour-optimised and "drawn" (made up) to conform to the "prudent" trend.

A further complication in applying Hood's typology to actual practice is related to the predominantly informal nature of incentives in the Russia civil service. When classifying

performance regimes into *target*, *ranking* and *intelligence* one cannot simply look at the officially declared mechanisms of linking performance to rewards and sanction. Informal rewards and sanctions have to be considered as well. If an indicator is officially regarded as an *intelligence* indicator in policy documents, it does not mean that it is regarded similarly by governors or other civil servants. Lower level civil servants may still engage in data manipulation driven by informal incentives to be recognized by their superiors.

It is hard to single out a particular example, because higher-than-average growth of an indicator does not automatically mean that this indicator is being manipulated. It should be noted, however, that, despite not being included in the formula for grants, all indicators related to sport demonstrated outstanding growth in the period between 2007 and 2011. In the run up to the 2014 winter Olympic games the sphere of sport received increased attention. This could create additional informal incentives to drive these indicators upwards. At least some portion of this effect may be the result of upward manipulation.

One of these indicators deserves particular attention: "The number of people taking part in organized sport events". One respondent from a Siberian region (lower-middle tier civil servant, *vedushiy specialist*) told the following story:

> *In our district we had to organize free outdoor aerobics sessions for citizens to attend. They were scheduled to take place each week on Tuesday at 2pm at our local stadium. Naturally, no people came. But they had to show to the journalists that the attendance was high, so they could include it into an official report and broadcast it in the news. So they commanded us (civil servants) to come to these 1-hour long aerobics sessions so that the journalists could film a crowd of sports enthusiasts. And we had to attend because they said that they would check attendance and reprimand those who would be absent.*

This example indicates that informal incentives could be as (if not more) powerful than formalised rewards. Besides, if it is accurate that at the federal level the system of indicators was chiefly used to provide the president with a bargaining advantage in his negotiations with

governors[52], then the distinction between "ranking" and "intelligence" indicators becomes even more blurred. The president could discretionary choose an indicator from the set of 325 titles and pick it to reproach the governor for poor performance. In this case all indicators become, in a sense, ranking indicators. So the difference between types of indicators becomes one of degree, not of kind. Indicators included in league tables are *more likely* to attract attention of the superior, but other indicator also may be under pressure to be manipulated.

Informal incentives may intensify the theoretically predicted effect of negative selection against honesty in bureaucratic hierarchies. This effect is discussed in the following section.

Performance reporting as "bureaucratic panopticon"

In his book *"Discipline and punish"* Michel Foucault used a metaphor of a "panopticon" to analyse the nature of power relations in modern society (Foucault 1977). Panopticon was a "perfect prison" conceived of by an 18th century English philosopher Jeremy Bentham. This prison was designed as a circular structure with a watch tower in the middle and cells arranged around it in such a way that at any given moment the watchman could observe the prisoners, but the prisoners were not able to tell whether they were being observed. A prisoner would have to constantly monitor his behaviour as if he was being constantly watched. The beauty of the system was in the fact that it made it unnecessary for there to be a watchman at all. Foucault used this metaphor to illustrate his notion of "normalization" – the process by which individuals in modern society are made to conform to dominant norms. Modern technology creates a situation in which individuals are aware of the possibility of being under constant surveillance and, thus, they internalize conformity.

[52] Refer to the discussion of the psychological effect of indicators in Chapter 5 (qualitative results). A federal civil servant with thorough experience of the system observed:

> *The only real reasonable function of this system is that it allows us to provide the president with an analytical brief when he is meeting a governor. So that the president can look at the table with performance figures and ask the governor "Why is the value of this indicator 5 in your region, while in other regions it is 8?". This is important, because it allows the discussion to begin. Then the governor has to respond and defend himself and so on. But functionally, there is nothing.*

This metaphor may be fruitfully applied to the system of performance reporting in Russia that seemed at the first glance an empty bureaucratic exercise. Regional governors were required to send their performance reports to the national government (Ministry of regional development and the presidential administration). They perceived the danger (real of illusory) of getting unwelcome attention if their performance figures deviated too strongly from the trend that they perceived as normal (or thought that the central government would perceive as normal). Both unexpected rapid growth and a sharp decline in performance figures could lead to unwanted scrutiny by the central authority. Thus, the most prudent way of keeping a low profile was to normalize performance figures to conform to the perceived normal trend even in the absence of formal instructions or incentives to do so.

A strong tradition of informal relationships within bureaucratic hierarchy in Russia could have magnified the effect of this normalization stimulus, but it is, I believe, common to any bureaucratic system of reporting. It is consistent with this interpretation that the system of reporting that existed in Russia at the time had very weak formal incentives to improve performance figures. Respondents were unable to give an example of formal sanctions for bad performance.

This interpretation of motives behind data manipulation links well with the model of "administrative man" developed by 1978 Nobel Prize laureate in Economics Herbert Simon. Among Simon's many contributions is the development of the theory of "bounded rationality". His "administrative man" (in contrast to the "economic man" of classical economic theory) has a limited ability to process information and a limited "focus of attention" (Simon 1954). For instance, a supervisor can at any one moment monitor only a small portion of his subordinates' activities. Being aware of this, the majority of subordinates try not to attract their superior's attention by behaving "normally" and staying "under the radar". It is only when some unusual activity is noticed does a supervisor begin to investigate more thoroughly and directs his focus of attention toward some particular activity.

I believe that by fusing Foucault's theory of normalization and Simon's model of administrative man one gets a productive interpretation of the phenomenon of "prudent" data manipulation. Civil servants in the process of formal performance reporting attempt to stay out of their superior's focus of attention by reporting "normal" performance data. The relationship between a reporting agent and his supervisor is somewhat akin to the relationship between Frodo and the Eye of Sauron in *The Lord of the Rings*. One can get away with practically anything as long as one does not attract the scrutinizing gaze.

This interpretation explains the presence of "prudent" strategy in the Russian case. In addition to negativity bias traditional to the public sector, this suggests a certain "attention aversion" that characterizes internal bureaucratic relations. It seems likely that such attitude is shared by many hierarchical systems, both public and private. Bureaucratic transparency may render effects of such attitude quantifiable. Efforts to be inconspicuous, if taken systematically by a large proportion of reporting agents, produce a distortion that becomes itself conspicuous.

Hermetic bureaucracy as a natural laboratory

The data provided by Russian national performance dataset, has some strong advantages from academic point of view. The data were collected by different agencies creating nearly "natural experiment" conditions and allowing to delineate the effect of varying types of principal-agent relations on reported data.

The system created favourable conditions for the observation of strategic behaviour of bureaucrats. Longenecker and Ludwig (1990:968) argue that, in order for fairness and accuracy to be expected from reporting agents, a performance measurement system has to satisfy a number of criteria. These criteria include *a sound procedure* of reporting that does not impose unnecessary costs on participating managers; adequate *training* ensuring that people understand how and why the system operates; *leadership* from above demonstrating that performance data are taken seriously; and an appropriate *auditing* procedure to insure overall integrity of the process. It became evident during the interviews that none of these conditions was met in the

implementation of this particular system[53]. Thus, the system may be said to have failed to create necessary conditions for fairness. This is a drawback from the point of view of organisational design, but an advantage from the point of view of data manipulation research as it made the occurrence of data manipulation practices more likely.

Another reason why the Russian case may be considered exemplary is the relatively high degree to which the Russian system of civil service is insulated from external scrutiny. The Russian "bureaucratic leviathan" (Shinar 2012) has a life of its own. Neither the civil society, nor politicians have been able to establish effective control over public sector bureaucrats. This is generally considered a great flaw of the Russian system of civil service, but it is an advantage for academic research of bureaucracy as it allows one to observe pure bureaucratic behaviour, untarnished by interventions of other systems.

If this interpretation fairly reflects the nature of the system, then performance data produced by it may be seen as largely symbolic and only indirectly linked to reality. The data certainly reflect something, but this something may be not organizational performance, but rather traces of bureaucratic survival strategies of individuals and organizations.

Characterizing Russian regional context

How, then, can we characterize organizational culture surrounding performance measurement in Russian regional governments? Based on our findings we can conclude that 1) performance management reforms were nominally an important part of administrative reform, but in practice "bobbed on the surface" of the organizational life; 2) the type of transparency introduced by the national system of performance measurement may be classified as "bureaucratic"; 3) the system combined *intelligence* and *ranking* regimes, but due to spontaneous informal incentives some indicators were treated as if they were *all ranking*; 3) in relation to

[53] Respondents were very critical about the implementation process, particularly the lack of feedback and consultations. One respondent remarked: "*[the authors of the reform] must have taken these indicators from Europe, changed them a bit to correspond to Russian reality and rolled them out. And that was it. Now we must work with them, whether we like them or not".*

performance measurement organizational culture may be classified as “fatalist”, indicators were imposed by centre and never became “owned” by regional civil servants. Regions minimized costs of compliance by insulating most departments from the routine of performance reporting and concentrating performance management staff within one department (frequently a department of the Ministry of economic development). Most staff were not involved in performance management exercises. Indicators were selected without explicit assessment of the type of function that were to be measured. Thus, no distinction between canonical and non-canonical practices was made. In many cases output measures were assigned to tasks that are hardly measurable (in health or sport).

Alienation of staff from the system of performance measurement was pronounced, feedback was non-existent and understanding of the baseline was low. The implementation of performance measurement was not accompanied by managerial devolution. In many instances it was the other way around. Standardization and administrative regulation of most practices in the last 10 years contributed to turning regional ministries from managers to administrators of federal programmes.

Administrative values and appropriate reform toolkit

In this section I attempt to fuse two typologies to explain the fact that performance measurement failed to take deep roots in the Russian public sector. First I look at Michael Barber’s typology of public sector reform styles that shows what reform instruments are appropriate for different stages of reforming public services. Second, I look at Christopher Hood’s typology of administrative values that orders different reforms according to a hierarchy of public sector values. By bringing together these two typologies I show that performance measurement could not take deep roots in the Russian public sector because it was implemented in the context where fundamental “capital base” of public sector values was not yet in place.

Barber (2007:238) identifies four levels of the quality of public services: awful, adequate, good, great. These four levels of quality correspond to four types of client attitudes: exiting,

grumbling, staying, committing (Figure 31). According to Barber, the aim of public policies is to increase the level of quality of civil services until they may be ranked as "great". This was the task of the Prime Minister Delivery Unit under Tony Blair. In order to achieve this change, it was important to engage both citizens and civil servants. It is futile to rely on citizen participation in improvement of "awful" services: clients try to avoid using such services. Similarly, it is useless to try to engage civil servants in service improvement if basic conditions of service delivery are not met. Barber suggested a typology of methods or public service reforms that correspond to different levels of quality of public services.

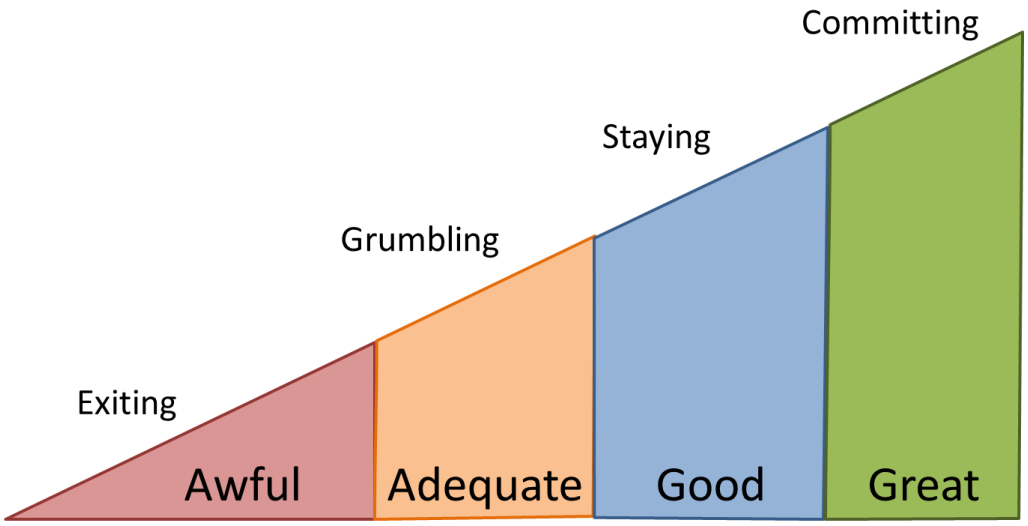


Figure 31. Public sentiment towards the quality of public services (Barber, 2007:238).

Typology of methods of reform

Barber identifies three methods of public sector reform: "command-and-control", "quasi-markets" and "devolution and transparency" (Figure 32):

Command-and control

The most effective method of bringing the quality of public services up from awful to adequate is through strict centralized control. This method allows for rapid improvement of failing services and brings noticeable results quickly. A typical example of the use of this

method is the military. Characteristics of this method are: regulation, personal responsibility of administrators and strict hierarchy. This may be called a "militarist" approach to reform. Elements of militarism allow for quick build-up of lambda-values of resilience and reliability. This method has its upper limit: it is impossible to reach "good" and "great" quality through direct command-and-control:

> *«…command-and-control cannot deliver 'good' or 'great' for… 'You cannot mandate greatness; it has to be unleashed.'»* Barber (2007:337).

This method of reform is unsustainable: as soon as pressure is removed, the service relapses. Sustained attention of the central authority is required to maintain the process of reform. It may require the use of "manual control" if some of the aspects of a service are failing. Agents have no incentives for initiative. Initiative comes from the top and agents are reduced to mere administrators. Agents' initiative is, however, a valuable resource that may be tapped into. It is made possible by the two other methods: "quasi-markets" and "devolution and transparency" (Barber, 2007:337).

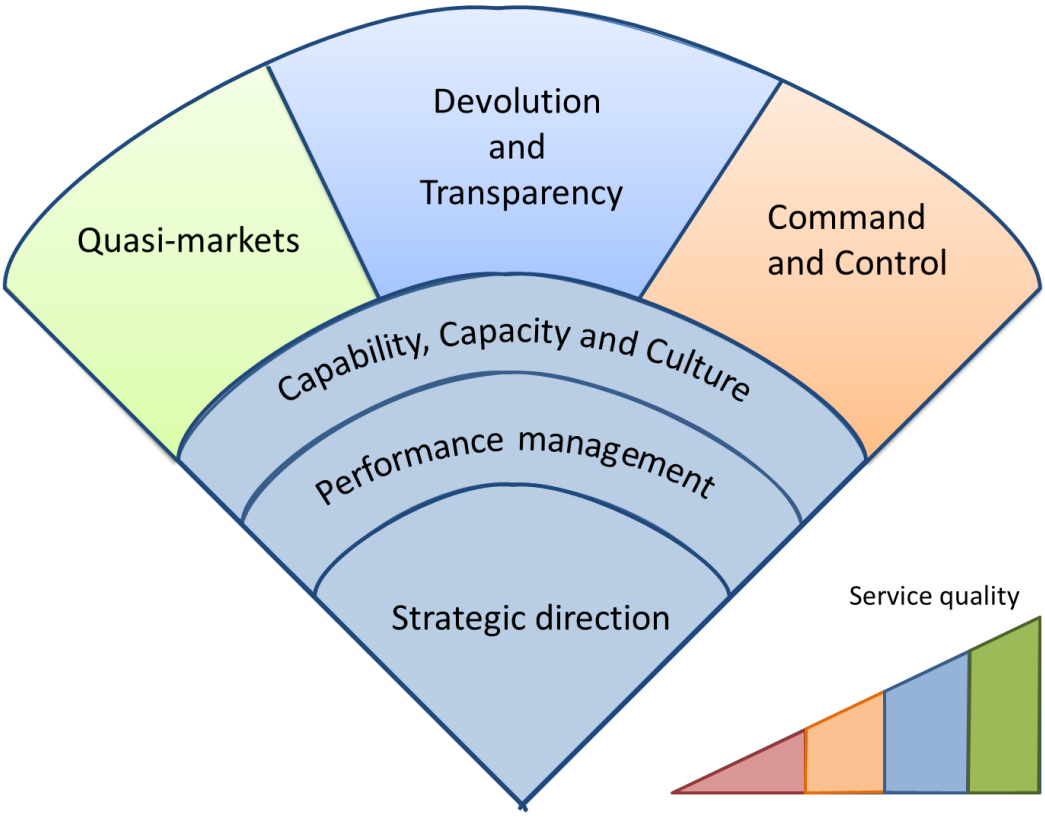


Figure 32. The typology of methods of public sector reform. Barber (2007:336)

Quasi-markets

The second method of reform that allows to unleash the energy of self-monitoring agents is termed by Barber as "quasi-markets". This method aims at initiating the process of improvement through competition among agents. The ultimate goal is to move from unsustainable enforced improvement to sustainable development:

> *Innovation should come from self-sustaining systems.*
>
> Tony Blair, cited in Barber (2007:337)

Quasi-markets may be created through giving greater freedom of choice to clients and stimulating competition among service providers. Reform may be driven by this method when clients stop avoiding the failing service and start actively using it because its quality becomes acceptable to them. When this happens, clients may be engaged in sustained service improvement. Similarly, competition among service providers is only possible when they are interested in attracting more clients. In the private sector profit provides the motivation for

service providers. In the public sector motivation may be created through the system of performance management: rewards, rankings and sanctions may be used to create quasi-markets. Barber lists a number of pre-requisites for this method: clients need to have choice, there should be independent service providers, there should be some variation in quality of services. Education and health services are examples. League tables are among the instruments of setting up quasi-markets, other instruments include educational vouchers, leading service providers may be rewarded and given greater managerial autonomy. Barber (2007:338) emphasizes the importance of government regulation to prevent undesirable effects such as growing differentiation among service providers

Devolution and transparency

The third method is termed "devolution and transparency". This method may be used when conditions do not allow for the establishment of quasi-markets. It may not be possible to allow clients to choose service providers (as in the case with policing). In this case competition may be created artificially through a number of performance measures and publication of rankings and league tables. Transparency is achieved through performance measurement. Good performance may be rewarded by greater devolution. This method combines centralized target-setting with devolved managerial discretion. According to Barber (2007:339) this method was effectively used in New York police and in British prisons.

Combining methods of reform

Barber (2007:339) notes that all three methods may be combined and used together. For example a school may be either subjected to direct command-and-control or devolution depending on its performance and "maturity". Parents may be given vouchers to allow them to choose schools for their children, while the school management may be given discretion to spend money with relative autonomy. At the same time, exam results may be published in league tables leading to rewards or sanctions. So all three methods may be combined.

Matching administrative values to methods of reform

The hierarchy of administrative values (Figure 8) (Hood, 1991) and the triad of reform methods (Figure 32) (Barber, 2007) may be matched (See matching colour on Figure 33).

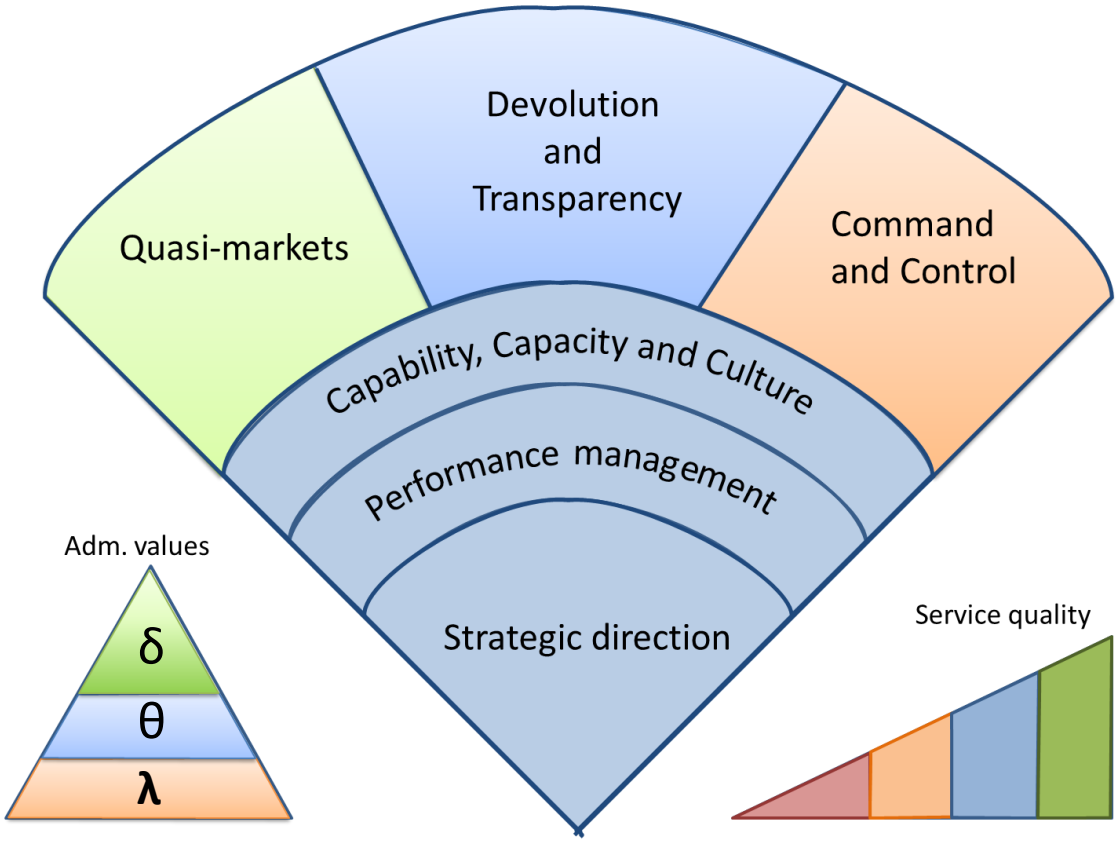


Figure 33. Matching reform methods and administrative values.

Command and control is effective in promoting lambda-values of resilience and reliability and promoting dependable service provision. If this is the aim, then considerations of economy and transparency play an insignificant role. When minimal conditions of a dependable service provision are met, further improvement becomes possible. It is futile to try to promote public sector ethics and transparency when a service is struggling to deliver altogether because of crises and failures. Only a working mechanism can be improved. Established lambda-values are a necessary but not sufficient condition for service improvement.

The second key condition for using more sophisticated methods of reform is the established base of theta-values: integrity, transparency, ethics and the rule of law. The two remaining methods require managerial discretion to be shifted away from the supervising

authority. Local managers should be given discretion over their budgets. It is assumed that established theta-values ensure that money is spent in the interest of the public. Increased autonomy would unleash local initiative and kick-start the process of self-improvement. If theta-values are not yet established, increased discretion may lead to adverse consequences: rent-seeking and corruption.

The two methods (quasi-markets and devolution and transparency) focus on promoting sigma-values of economy, efficiency and effectiveness, while taking the capital base of lambda- and theta-values for granted. Both methods rely on competition between service providers to ensure sustained improvement. Hood (1991) warns that over-reliance on competition may undermine the capital base of traditional values. Excessive borrowing of private sector management practices may erode traditional values of "militarized" bureaucracy and render it incapable of performing its tasks. Unintended consequences may cause overall performance to drop.

In the study respondents who had first-hand experience with designing the performance measurement reform package a number of times used the expression "*navodit' poryadok*" (to sort things out, to establish basic order) referring to the original aim of the performance measurement movement. This is clearly a case of attempting to build lambda-values with a command-and-control methods to bring service provision from awful to adequate. This may partially explain why from the initial package of the Administrative reform most significant amount of resources was spent on "administrative regulations" - standardization and procedure-control tools that assigned formal descriptions to public services. Managerial devolution (a tool of promoting sigma-values), never appeared on official agenda. The time was right for improvement of lambda-values of basic sustainability, reliance and resilience. Sigma-values of economy and local initiative were premature.

Administrative values in the Russian case

> *The lack of transparency is perhaps the most important legacy of the Soviet era that continues to beguile the Russian public service and its reformers.(Barabashev and Straussman 2007: 376)*

In the Russian case it appears that public sector reforms have not yet reached the stage where measures promoting sigma-values are timely. A number of times top-level respondents from among those with first-hand experience of designing the reforms mentioned that the chief aim of reforms was to ensure basic order in inter-governmental relations (*navesti poryadok*). It seems that this was the aim when first DRONDs were implemented in 2004. According to Hood's scheme it is logical that no devolution was used at this stage as more "militaristic" measures were more appropriate.

Perhaps, performance measurement in Russia was implemented ahead of its time. Currently, it seems that most attention of public sector managers is drawn to the transparency of the reforms (Open Government, open data, transparent budgets, etc.). These are initiatives aimed at building theta-values according to Hood's scheme. Perhaps, later the time will come for a sigma-values reforms such as managerial devolution accompanied by robust performance management regimes. However, it may take a lot of efforts to change the existing transparency-averse culture of the public sector largely because of Soviet legacy:

> *Soviet bureaucrat's job was to implement and even predict the leadership's demands and to follow the party guidelines, but never to work directly in the interest of citizens — a bourgeois concept, at best. This lack of transparency, endemic to Leninist systems, carried over to Russia after 1991 and has been slow to change. (Barabashev and Straussman 2007: 376)*

In contrast, reforms aimed at promoting lambda-values of reliability and dependability of services seem to take root in the Russian environment. This focus on essential functions of governance is necessary to rebuild the public sector after the collapse of the Soviet regime:

> *Any post-Soviet strategy of political- economic development . . . would have to aim as a necessary precondition of success at establishing and maintaining a system of public administration, including a civil service, that would enable the state to perform the*

*essential functions of governance (Lynch 2005, 83 cited in Barabashev and Straussman (2007: 381)***).**

Administrative values: deficit of theta-values

The way performance indicators came to be used in the Russian public sector may be interpreted as a manifestation of the fact that the "capital base" of lambda and theta values was insufficiently developed. The atmosphere of public sector honesty is a prerequisite for reforms addressing issues of efficiency and effectiveness. It is hard to muster solid evidence of this fact, but I believe that the sum of evidence presented in this work provides at least some support for this assertion. The way performance indicators are used in the Russian public sector suggests that they have been implemented in an environment where honesty and desire to serve the public good is still yet to be developed. Further evidence of this fact may be discerned in the results of the realization of 2006-2011 administrative reform. As outlined above (see above: Administrative reform package: contradictory intentions) the reform package included a wide variety of elements. "Administrative regulation" was the element that received most publicity and awareness. This part of reform package aimed at establishing a rigorous process of control over public sector organisations. Administrative regulation and standardization remain the *leitmotiv* of modern reform efforts. Needless to say, thorough standardization and process-control are at odds with managerial autonomy and devolution of managerial discretion. These latter elements are integral part of western-style performance management. Therefore, there is a inherent contradiction between real performance management and real top-down standardization. This contradiction is not unique to Russia.

Hood (1995) observed that, once fully implemented, the ideas of NPM, including performance management and managerialism, manifest their weaknesses. As with many other things, it appears that in this case, too, the best way of undoing is overdoing. Unchallenged ideological dominance of NPM has in many cases brought about "the self-censorship effects of [its] cultural hegemony: compliance, caution and risk avoidance in the interests of survival." (Barrett, 2004, p.259).

On the practical side, the "mixed-salad" approach of NPM, whereby managerial techniques borrowed from different frameworks are combined to provide the synergy and offset weaknesses of each technique, results in "mutual repulsion" between the elements (Hood, 1995, p.110). Indeed, in discussing NPM one frequently sees mutually exclusive concepts blended together – decentralisation and centrally imposed targets, coordination and competition, power to manage and accountability.

Moreover, when thoroughly followed to their extremes, NPM policy recipes tend to turn into "fatal remedies" (Hood, 1995, p.112). Side-effects accumulate and unintended consequences manifest themselves strongly.

It appears that in the Russian case centralisation, standardization and process control ended up making most of the "salad" with only occasional and decorative sprinkles of genuine performance management. Negative selection (both entry-level and through promotions in the hierarchy) may have contributed to the erosion of theta-values in the Russian public sector. If this has been the case, then performance management reforms could not avoid the deforming pressure. They had to be sculptured to fit into the framework of military-style centralised bureaucracy with the dominance of self-interested motives, compliance culture and risk-aversion. Instead of helping to change these features PM may have contributed to their entrenchment. Formal decorative use of proper lexicon provided by PM could help disguise the reality of rent-seeking bureaucracy.

Performance indicators as a tool of cooperation

> *It is difficult, sometimes impossible, for one man to coordinate his own actions so that these lead to the accomplishment of a desired goal. Such coordination should not be expected from a group composed of several separate individuals. … Coordination will always be much less perfect. Administrative organizations cannot be assigned tasks that require perfect coordination, or even some approximation of this (Tullock 2005: 163).*

Many of the aims that public sector organisations work to achieve are a result of the work of multiple agencies. Establishing the contribution of an individual organisation may not be possible or feasible. If an aim can only be achieved by multiple agencies in cooperation, they need to coordinate their efforts. Performance management systems can be instrumental in promoting both vertical and horizontal cooperation. Vertical (centralized) cooperation may be promoted through the system of cascading objectives that enable many organisations to work on achieving a shared higher level objective. Potentially, they can also promote horizontal cooperation by allowing different agencies to "speak the same language" and monitor cooperative efforts.

Horizontal cooperation - is one of the most widely discussed topics in modern studies of government: New public management (D. Osborne and Gaebler 1993) is being replaced by New public governance (S. P. Osborne 2010) and Network governance (Bogason and Zølner 2007; C. Jones et al. 1997; O'Toole 1997).

Imperial (2005) notes that the task of facilitating interagency cooperation is a difficult one and requires a comprehensive approach. The likelihood of successful cooperation is influenced by such factors as the personality of the leader, the procedures of budgeting, planning, human resource management and procurement. Particular importance is attached to the degree of trust between the partners: "cooperation - is an advanced management method "(Imperial 2005: 308).

One Important factor of successful cooperation is more or less even distribution of resources between organizations. If resources are distributed without obvious imbalances, the conditions for horizontal cooperation are favourable. If resources are concentrated in a single organization, its bargaining power is disproportionate, and it is difficult for other partners to ensure a "win-win" mode of cooperation (Imperial 2005: 309).

Cooperation can be achieved by the means of either centralized or decentralized mechanisms (Bardach 2001). Recent studies show that cooperation on the basis of decentralized

mechanisms (networks, horizontal interactions) may be more effective and sustainable than forced cooperation, imposed from above (Imperial 2005).

Mechanisms of decentralized (or polycentric) cooperation have received significant attention of scholars in recent years, as evidenced by the award of the 2009 Nobel Prize in economics to Elinor Ostrom, whose research focused on problems of collective action and horizontal cooperation.

Imperial (2005) has shown that, although centralized coordination creates more transparent control mechanisms, it can be inferior to decentralized coordination in effectiveness. Transaction costs of using a central coordinating agency may be higher, and a number of indirect benefits from cooperation are foregone. Such benefits include knowledge sharing, generation of social capital, increased trust between partners, strengthened morale and greater initiative. In addition, horizontal cooperation can promote innovation and facilitate dissemination of best practices (Ostrom 1990, 1998; Ostrom et al. 1999; Rogers 1996).

But, perhaps, the most important conclusion that can be drawn from the literature on polycentric cooperation is that forced cooperation under central supervision is not sustainable. Once central supervision is withdrawn, cooperation ceases. Horizontal network cooperation, on the other hand, may become self-sustaining if certain conditions are present (Dietz et al. 2003; Kuzminov and Yudkevich 2010; Ostrom 1999).

Horizontal cooperation provides an alternative mechanism for collective action, but its applicability is limited to areas where there is a possibility of mutual benefit. In cases where there is a conflict of interest, centralized coordination may be more appropriate.

Interdepartmental interaction fits into the framework of the theory of collective action (Hardin 1968; Olson 1971; Ostrom 1999). A situation in which agencies work together to improve one or several performance indicators is similar to that in which individuals embark on a collective undertaking. Costs are incurred by one institution but benefits are shared by all, so that a free-rider problem arises.

Olson (1971) argues that a collective action will not take place unless participants take personalized interest in the success of the action: “[o]nly a separate and “selective” incentive will stimulate a rational individual in a latent group to act in a group oriented way” (Olson, 1971, p.51). That is, a mere fact of promoting the "common good" is not sufficient to motivate a rational member of a group to actively engage. There must be selective benefits or coercion. Similar problems arise when resources are in common use (see “The tragedy of the commons” (Dietz et al. 2003; Hardin 1968; Ostrom 1998, 1999, 2000).

The table below summarizes “design principle” of successful cooperation according to Ostrom (2000).

Table 24. Design principles of successful cooperation.

Design principle	Original meaning (Ostrom, 2000)
The presence of clear boundary rules	This principle enables participants to know who is in and who is out of a defined set of relationships and thus with whom to cooperate.
Allocate benefits proportional to required inputs	The second design principle is that the local rules-in-use restrict the amount, timing, and technology of harvesting the resource; allocate benefits proportional to required inputs; and are crafted to take local conditions into account. …How to relate user inputs to the benefits they obtain is a crucial element of establishing a fair system.
Most of the individuals affected by a resource regime can participate in making and modifying their rules.	Resource regimes that use this principle are both able to tailor better rules to local circumstances and to devise rules that are considered fair by participants.
Select monitors, who are accountable to the users or are users themselves	Most long-surviving resource regimes select their own monitors, who are accountable to the users or are users themselves and who keep an eye on resource conditions as well as on user behaviour.
Graduated sanctions that depend on the seriousness and context of the offense	In a regime that uses graduated punishments, however, a person who purposely or by error breaks a rule is notified that others notice the infraction (thereby increasing the individual's confidence that others would also be caught). Repeated infractions attract more and more serious sanctions. The capability to escalate sanctions enables such a regime to warn members that if they do not conform they will have to pay ever-higher sanctions and may eventually be forced to leave the community.
Cost-effective local conflict resolution mechanisms	By devising simple, local mechanisms to get conflicts aired immediately and resolutions that are generally known in the community, the number of conflicts that reduce trust can be reduced.
Minimum recognition of the right to self-organization by the authorities	This allows local users to create rules that match their local conditions, reduces risk of interference from superior authorities and legitimizes local procedures.

The importance of feedback

As qualitative findings suggest, in the Russian case insufficient attention was paid to securing the "buy-in" of public sector staff. The importance of employee resistance in the public sector should not be underestimated. From the very top to the very bottom of the bureaucratic hierarchy, implementation of any task depends upon the will of the public sector workers. This was true a century ago, this is true today:

> *The Russian czar of the old regime was seldom able to accomplish permanently anything that displeased his bureaucracy and hurt the power interests of the bureaucrats (Weber, 1948/1922, p.238).*

In summarizing major lessons learned from implementing performance management in public and private sector Likierman (1993) placed the importance of feedback among 20 important lessons (lesson 5), and cited two of his respondents to illustrate his point:

> *'Attitudes change depending on whether you are calling for performance indicators or you feel they are being imposed on you' commented the finance officer of a quango, and a senior executive of a central government agency emphasized that: 'Agreeing performance indicators is a negotiating process in the broadest sense. If this is ignored it will lead to poor commitment and sense of ownership', and went on 'People must understand what is expected of them, and how this was decided. They must be allowed to contribute to the decision-making process'. (Likierman 1993: 17)*

CHAPTER 8. ALTERNATIVE INTERPRETATIONS OF THE ROLE OF PERFORMANCE MANAGEMENT

A note on understanding as an epistemic tool

This section offers a discussion of the method of *understanding* as an epistemic tool of this research project. This method has been used throughout this research project. I cite two prominent thinkers who advocated the use of this method in social sciences: Gordon Tullock and Ludwig von Mises. It is worth citing Tullock's discussion of the method of "understanding" at length:

> *We have, basically, three ways of finding things out. First there is mathematics, or pure abstract thought. In a sense this is an exploration of the logical categories of the human mind. A second method is observation of our environment, a category which includes the "highly scientific" processes of experiment and investigation carried on in laboratories. The third method of finding things out I should like to call, with Max Weber, "understanding." In a sense this is as introspective as mathematics. We understand how others feel or act, because we know how we would act or feel under similar circumstances. This method, used by practically everyone in everyday life, is not applicable to the physical sciences for obvious reasons. In investigating die properties of a chemical, little progress can be made by saying to yourself: "If I were sodium hydroxide, what would I do?" <….> [The] principal sphere of usefulness [of understanding] is obviously in the study of human beings. In recent years students of human phenomena have sometimes tried to avoid the use of this tool. This appears to spring from a misunderstanding of the situation. The physical scientists, and particularly the physicists, have established positions of great prestige in the present-day learned world. For the reasons given above, they make almost no use of "understanding" in their work. From this a number of "social scientists," anxious to establish their claim to be real scientists, have deduced that this method is "unscientific" and to be avoided.*
>
> *Except for the questionable purpose of gaining social prestige within a university faculty, however, there seems to be no good reason for deliberately refusing to employ this method of investigating human behaviour. The fact that the social scientist can use this tool which is unavailable to the physical scientist should be considered an advantage. This is not to suggest the abandonment of other methods. The problems*

are difficult, and the discarding of any tool that may assist in their solution would be unwise. (Tullock 2005: 17-18)(underline added)

The major implication of the use of this method is that it opens up the possibility for the usage of an additional type of evidence – the intuition and experience of the reader. Tullock explains this in the preface to his major work "Bureaucracy":

The use of this approach or method has, however, a strange, even paradoxical, consequence. For a number of the assertions that will be made in this book, the supporting evidence must be found in the mind of the reader. That is to say, instead of presenting concrete evidence, I shall simply try to convince the sceptical reader by appealing to his own intuition and experience. I shall offer examples of [selected] types of behaviour, not to prove particular points, but simply to explain the points better to the reader so that he may judge whether human beings, in general, behave in the manner that I suggest. If his "understanding" leads him to the same conclusion that I have reached in a particular case, he can then accept my statement of the principle as being true (Tullock 2005: 18)

Another strong defence of the use of understanding as a distinct method of social sciences was advanced by an Austrian economist Ludwig von Mises. According to Mises, this "understanding" is not some mysterious wisdom possessed by professional academic scholars. Understanding as a scientific tool is applied by historians, but it is also the tool every laymen use in their everyday life:

The intellectual methods of science do not differ in kind from those applied by the common man in his daily mundane reasoning. The scientist uses the same tools which the layman uses; he merely uses them more skilfully and cautiously. Understanding is not a privilege of the historians. It is everybody's business. In observing the conditions of his environment everybody is a historian. Everybody uses understanding in dealing with the uncertainty of future events to which he must adjust his own actions. The distinctive reasoning of the speculator is an understanding of the relevance of the various factors determining future events. <…> [A]ction necessarily always aims at future and therefore uncertain conditions and thus is always speculation. Acting man looks, as it were, with the eyes of a historian into the future. (Mises 1949: 58)

This concluding remark that "acting man looks with the eyes of a historian into the future" tells us much about the nature of understanding as an epistemic method. It is the procedure through which we assign relevance to different factors that are likely to contribute to the outcome of a situation. In case of history – it is the evaluation of relevance of difference factors that contributed to an outcome of an event in the past. Since in social sciences we are always dealing with complex phenomena and human actions, no mechanistic algorithm can do the job of understanding for us. Quantitative assessment of the "magnitude of coefficients" and their "statistical significance" is unable to substitute for the evaluation of relevance of different factors provided by understanding:

> *Experience concerning human action differs from that concerning natural phenomena in that it requires and presupposes praxeological knowledge. This is why the methods of the natural sciences are inappropriate for the study of praxeology, economics, and history (Mises 1949: 40)*

This statement was echoed by Tullock when he remarked that social sciences were different from natural sciences because they could make use of an additional source of information. A researcher can ask himself – "What would I do if I was in this situation?". A case study is then a method of acquiring sufficiently detailed description of "this situation" and positions of actors in it. The description should be thick enough to allow the researcher to use his praxeological reasoning and empathy to put himself, as it were, in the social actors' shoes and make judgments concerning relevance of different factors in determining the outcome of a situation in question.

In other words, in order to understand why such and such effects took place when such and such intervention was made, quantitative models are inadequate. What is measured and statistically analysed is always historical data relating to an unrepeatable historical event. It does not provide predictions:

> *The impracticability of measurement is not due to the lack of technical methods for the establishment of measure. It is due to the absence of constant relations. If it were only caused by technical insufficiency, at least an approximate estimation would be possible in some cases. But the main fact is that there are no constant relations. Economics is not, as ignorant positivists repeat again and again, backward because it is not "quantitative." It is not quantitative and does not measure because there are no constants. Statistical figures referring to economic events are historical data. They tell us what happened in a non-repeatable historical case. Physical events can be interpreted on the ground of our knowledge concerning constant relations established by experiments. Historical events are not open to such an interpretation. (Mises 1949: 56)*

It should be noted that in the study of organisations, when using quantitative data, we always deal with historical data. If Mises was writing about economic history, then we ought to write about "organizational history". Mises differentiates between three types of knowledge: "apodictic certainty (the realm of praxeology), class probability (the realm of the natural sciences), and case probability (the realm of history)" (Mises 1949: xvi). In this chapter we are concerned with the latter.

It was an unrepeatable historical event that a system of performance indicators with certain characteristics was introduced in a particular environment of the Russian public sector in 2007-2011. The difference between class probability and case probability is that in the later the judgments on the likelihood of various outcomes are not based on judgments of quantity (statistical significance, magnitude, etc.), but on judgments of relevance. In attempting to understand the fate of the system of performance indicators we must assume the role of historians of the public sector. And, as historians, the main epistemological instrument for us has to be understanding:

> *The historian can enumerate all the factors which cooperated in bringing about a known effect and all the factors which worked against them and may have resulted in delaying and mitigating the final outcome. But he cannot coordinate, except by understanding, the various causative factors in a quantitative way to the effects*

> *produced. He cannot, except by understanding, assign to each of factors its role in producing the effect. Understanding is in the realm of history the equivalent, as it were, of quantitative analysis and measurement. (Mises 1949: 56)*

Previous chapters have already covered a number of factors that, according to my judgment, could contribute in the degeneration of the system of performance measurement in the Russian case. It is, however, open for debate whether the factors covered were the most relevant. Historical investigation is open to such debates because of its subjective element:

> *To every historical factor understanding tries to assign its relevance. In the exercise of understanding there is no room for arbitrariness and capriciousness. The freedom of the historian is limited by his endeavour to provide a satisfactory explanation of reality. His guiding star must be the search for truth. But there necessarily enters into understanding an element of subjectivity.* ***The understanding of the historian is always tinged with the marks of his personality. It reflects the mind of its author.*** *(Mises 1949: 57)(emphasis added).*

For these reason, an explanation offered by one historian may be challenged by another:

> *Historical understanding can never produce results which must be accepted by all men. Two historians who fully agree with regard to the teachings of the non-historical sciences and with regard to the establishment of the facts as far they can be established without recourse to the understanding of relevance, may disagree in their understanding of the relevance of these facts. They may fully agree in establishing that the factors a, b, and c worked together in producing the effect P; nonetheless they can widely disagree with regard to the relevance of the respective contributions of a, b, and c to the final outcome. As far as understanding aims at assigning its relevance to each factor, it is open to the influence of subjective judgments. Of course, these are not judgments of value, they do not express preferences of the historian. They arc judgments of relevance (Mises 1949: 57).*

Then the persuasiveness of a narrative constructed by a researcher should be evaluated by the reader on the basis of his own intuition and experience. This is the ultimate measure of validity of a study of this kind.

Another consequence of the fact that in dealing with human action we are dealing with judgements of relevance is that a study of human action has to be holistic. It should attempt to include as much relevant context as its scope and human capacity permits. In this it is different from natural sciences. A study cannot be mechanistically limited to one aspect of a phenomenon, but has to select these aspects on the basis of relevance and practicality. Mises address the following quote to economists, but it equally applicable to students of organisations:

> *The empirical sciences start from singular events and proceed from the unique and individual to the more universal. Their treatment is subject to specialization. They can deal with segments without paying attention to the whole field. The economist must never be a specialist. In dealing with any problem he must always fix his glance upon the whole system (Mises 1949: 69).*

In our case these means that not only managerial environment in the public sector at the regional level is relevant to understand the fate of the system of performance measurement. Political and social environment in Russia at the time and in the past may also contribute to understanding the phenomena. The breadth of the study is limited only by practicality and judgment of relevance.

This is not to say that insights produced by such a study are not relevant for studies of other context. This only means that they are not generalizable in statistical sense (as when one speaks about generalizing from a sample to population), but they are generalizable using the same procedure of understanding. A researcher interested in lessons from the study of a foreign case must make a judgment regarding relevance of the findings for his own case. In doing so he should evaluate not only factors included in the original study by the author but also other factors that he himself finds relevant. It is possible that some of the insights (should there be any) produced by this study will be considered relevant by a wider community of scholars for a wide variety of cases, other findings would be considered relevant for a more narrow range of cases.

Additional support for the use of such method may be given by citing numerous authors who have recently turned to qualitative and interpretivist research in the literature on organization theory and other social sciences.

Having stated these considerations I now turn to a factor that I consider relevant in understanding the peculiar nature of the national system of performance measurement in Russia. I attempt to look at the practice of measuring performance as one of many modern ideologies of the public sector.

I explain the fact that performance measurement initiatives persist despite apparent universal scepticism by using Zizek's theory of ideology. Performance measurement may be interpreted as one of modern ideologies of the public sector that exist for the sole purpose of ensuring that we continue doing things in a certain consistent way, no matter how dubious and contingent our choice of this way may have been in the first place.

Performance measurement as ideology

Slavoj Zizek's theory of ideology can be usefully applied to interpret the findings of this project. My findings show that respondents were sceptical (and cynical) regarding the usefulness of performance measurement. Performance reports are seen as "documents for documents' sake". Despite this fact, the performance measurement system continues to exist and considerable resources of time, money and efforts are spent on maintaining it. In this section I argue that there is one particular function of the system that seems to be contributing to its continued existence despite this universal scepticism. Using Zizek's theory of ideology I interpret performance reporting as one among many modern ideologies of the public sector. In the same way as in the Soviet Union a Soviet official had to pledge his allegiance to ideals of communism, modern Russian bureaucrats pledge allegiance to ideals of "performance management", "New public management" and other managerial ideas. One cannot become a senior civil servant unless he embraces the importance of demonstrating positive growth of performance indicators, however meaningless and biased he himself may find them.

Despite universal scepticism (and cynicism) among civil servants regarding managerial usefulness of performance indicators, performance measurement procedures continue to exist and people continue to spend their time and efforts doing them. One regional minister of economic development told me that he thought that performance reports were just "documents for documents' sake". They were not linked to budgeting or any other practical decision-making mechanisms. Most respondents thought that performance data were not useful, yet they continued to collect and report them. Some respondents openly admitted that they themselves manipulated the figures or knew that others were doing it. Respondents were cynical about the truthfulness of official statements about effectiveness of performance management. I think that Zizek's theory of ideology can be applied to explain this attitude:

> *Cynical distance is just one way – one of many ways – to blind ourselves to the structuring power of ideological fantasy: even if we do not take things seriously, even if we keep an ironical distance, we are still doing them (Zizek 1989: 33).*

In Russia there is little evidence that the introduction of performance management procedures resulted in tangible improvement in public sector performance. Mostly these procedures are described by civil servants as meaningless cumbersome red-tape. Despite this, the central government insists on introducing more and more performance reporting procedures. The imperative to implement performance indicators persists in official government programmes without any justification or reference to assessment of the effects of previous reforms in this direction.

In attempting to explain this fact, one can use different perspectives. For example, it may be argued that civil servants are not supposed to understand the real usefulness of the system that they run. Maybe the system is considered useful by politicians? They may see some political purposes behind official empty policy statements. Or, it may be argued, the system is in place because of vested interests. An explanation may be constructed making use of such concepts as "regional elites", "bureaucratic elites", "consultocrats", "kleptocrats" etc. Indeed, consulting companies benefit significantly from dubious practices of designing systems of performance

measurement for governments. Corrupt civil servants may benefit from expensive management consulting contracts that brandish the label of "performance management" to cover up kick-backs and bribes. Another narrative may be constructed using such notions as "transparency", "accountability to the public", etc. Perhaps, performance indicators exist because they help civil activists to monitor public authorities. Some may interpret them as instruments of centralisation, others of decentralisation, etc. Success stories may probably be found where indicators were actively used to improve efficiency, and other failure stories may be found to demonstrate that they degenerated into empty red-tape and were never taken seriously. In still other cases they could have no effect on the practice of administration whatsoever. Each of these streams of reasoning has its own corresponding body of literature attempting to explain the role of performance management in various other practices and spheres of action. It would, probably, not be an exaggeration to say that a different fashionable managerial idea could be equally well substituted for "performance management" and would generate a similar host of perspectives, explanations and interpretations (it appears that currently in Russia "transparency" and "crowd-sourcing" seem to be taking the place of "performance measurement" as most fashionable buzzwords in policy documents and speeches of reformers of the public sector).

In this chapter I look at the brand "performance management" (together with such terms as "effectiveness of budgetary expenditure" and "strategic planning") as ideological signifiers. Performance management/measurement in the public sector in Russia have become a sort of ideology and has been persistently implemented with no attempts to provide factual evidence as a justification of their usefulness. Such treatment of a concept is a characteristic property of ideology. According to Zizek, ideology subsists only for the sake of ensuring that we continue doing things without questioning them (in the quote below he cites the metaphor of "travellers lost in a forest" by Descartes):

> *What is really at stake in ideology is its form, the fact that we continue to walk as straight as we can in one direction, that we follow even the most dubious opinions once*

> *our mind has been made up regarding them; but this ideological attitude can be achieved only as a 'state that is essentially by-product': the ideological subjects, 'travellers lost in a forest', must conceal from themselves the fact that 'it was possibly chance alone that first determined them in their choice'; they must believe that their decision is well founded, that it will lead to their Goal (Zizek 1989: 84).*

Performance management, as well as many other similar technocratic initiatives, may be perpetuated without an objective justification. Continued efforts at maintaining it are justified by reference to other signifiers such as 'effectiveness of budgetary expenditure' or 'strategic planning'. The materiality of these concepts is in itself dubious. They may be viewed as 'floating signifiers':

> *Ideological space is made of non-bound, non-tied elements, 'floating signifiers', whose very identity is 'open', overdetermined by their articulation in a chain with other elements… (Zizek 1989: 87)*

It may be that the whole structure of these ideological signifiers is held together by nothing substantial, only by a certain lack – the lack of purpose in actions of government bureaucracy other than these actions themselves. Most official documents list greater "effectiveness of budgetary expenditure" as the main aim of all reforms related to performance management. It is this signifier that seems to be performing the act of 'quilting' in major official documents:

> *The 'quilting' performs the totalisation by means of which this free floating of ideological elements is halted, fixed – that is to say, by means of which they become parts of the structured network of meaning (ibid).*

It is, however, perfectly consistent with Zizekian interpretation that it is never articulated what this expenditure is supposed to be effective at achieving. This "effectiveness of budgetary expenditure" is acting as a 'signifier without the signified':

> *The element which only holds the place of a certain lack, which is in its bodily presence nothing but an embodiment of certain lack, is perceived as a point of extreme plenitude (Zizek 1989: 99).*

From this perspective performance reporting mechanism can be seen as an exercise in evaluating 'ideological loyalty' of civil servants. In the same way as in Stalinist era a Soviet official had to pledge his allegiance to ideals of communism, modern Russian bureaucrats pledge allegiance to ideals of "effective budget expenditure" and all connected signifiers, such as "performance management" and "strategic planning". One cannot become a senior civil servant unless he embraces the importance of demonstrating positive growth of performance indicators, however meaningless and biased he himself may find them. Indeed, performance management may be a "useful myth"(Pollitt 2001) that allows one to separate loyal members of the group from outsiders:

> *Myth is not something simply 'untrue', an unexamined doctrine that will fall away in the face of rational thought and empirical evidence. Myth is the sometimes unstated and taken-for-granted 'common sense' of a policy community and as such, once established, is resistant to challenge.(Goldfinch and Wallis 2010: 1099)*

> *Adopting the rhetoric and structures and the legitimating version of 'reality' of this dominant myth … shows one to be of the 'in-group' of policy reformers with greater potential access to the power, consultancies and financial benefits that insider status entails. This is in contrast to the 'out group' that do not accept the myth; or are at least not clever or disingenuous enough to adopt its rhetoric, focus and belief systems – and who can as a result be portrayed and marginalized as intellectually moribund, out of date, recalcitrant, and/or having vested or rent seeking interests (Goldfinch and Wallis 2010: 1099).*

Moreover, established rhetoric excludes other potential discourses. By this I mean that habits of thought, speech and writing that surround the practice of performance management form a certain horizon of possibility for discussions. It is common to investigate the role of performance management in improving public sector efficiency, but one would hardly ever come across a discussion of the role of performance indicators in exploitation of front-line civil servants by senior bureaucrats. Such themes are excluded not only from official policy discourse, but also from scholarly discussions.

An example that has already been mentioned in the previous chapter may be a good illustration.

One respondent from a Siberian region (lower-middle tier civil servant, *vedushiy specialist*) told the following story:

> *In our district we had to organize free outdoor aerobics sessions for citizens to attend. They were scheduled to take place each week on Tuesday at 2pm at our local stadium. Naturally, no people came. But they had to show to the journalists that the attendance was high, so they could include it into an official report and broadcast it in the news. So they commanded us (civil servants) to come to these 1-hour long aerobics sessions so the journalists could film a crowd of sports enthusiasts. And we had to attend because they said that they would check attendance and reprimand those who would be absent.*

A more or less traditional interpretation of this episode would go along the lines of classifying this as an example of gaming, where civil servants distort their practice to meet targets. When discussed in such terms, it sounds more or less neutral. However, an alternative interpretation could focus on the fact that this episode was an example of direct violation of workplace ethics and, plainly speaking, a case of exploitation.

It is interesting to consider other instances of similar nature reported by the same respondent:

> *Respondent: We had to organize* subbotniki (*community environment cleaning day). We had to attend because they told us that we would have problems if we didn't come. They checked with a register. I didn't want to go, but they said that I would risk being fired. They provided no equipment, so we had to bring our own shovels and rakes. My friend accidentally pierced her finger with a rusty nail, and asked if she could go to the local medical centre to get it treated. The official who was in charge did not allow it, but she said she would go to court if he does not let her go. He threatened her that if she goes to court, she would not be able to prove that she worked involuntarily. Then she went to the hospital and came back to work an hour later than everyone else. They threated to fire her because she was not meeting the requirements of work discipline. She went to the HR office (*otdel kadrov*) and wanted to get a written confirmation that she was made to work at this cleaning event but they reproached her for her inadequate behaviour and said that they would not give any written references.*
>
> *Interviewer: Why didn't she leave this job?*

Respondent: Well, she's in her 50-s and she's always worked as municipal servant. What job could she get? She is afraid to lose her job.

Clearly, if such slave-like conditions are widespread (and they seem to be) at lower levels of the bureaucratic hierarchy, then discussions of the relation between "performance indicators" and "efficiency", "accountability", "transparency", "interagency cooperation" or "strategic planning" in Russia bear little in common with discussions of similarly sounding concepts in the USA or Britain. It may be that problems that dominate Western discussions of public sector effectiveness have very little relevance for Russia where such basic violations of human rights are common. The use of the same set of terms conceals underlying differences. And it also forces the discourse in Russia along Western tracks. In order not to sound backward and barbaric, people prefer to discuss such niceties as "performance management regimes", "open innovations", "crowdsourcing" and not plain and down-to-earth issues of exploitation, violation of human rights, poverty and misery. For, truly, "the limits or my language are the limits of my world".

It is worth recalling Max Weber's words on the importance of material dependency of civil servants on their organisations:

The individual bureaucrat cannot squirm out of the apparatus in which he is harnessed. … [He] is chained to his activity by his entire material and ideal existence. In the great majority of cases, he is only a single cog in an ever-moving mechanism which prescribes to him an essentially fixed route to march (Weber, 1948/1922, p.228).

Since Weber's times things have, undoubtedly, changed, albeit to varying degree in B. Material dependence upon the employer has fundamental consequences upon the performance of an employee. The degree of control that can be exercised upon individual behaviour of employees is negatively linked to the degree of their dependence upon the employer:

Other circumstances being equal, only economically independent officials, that is, officials who belong to the propertied strata, can permit themselves to risk the loss of their offices. Today as always, the recruitment of officials from among propertyless strata increases the power of the rulers (Weber, 1948/1922, p.235).

It is necessary, thus, to evaluate the financial positions of bureaucrats in both countries to understand the effects that such reforms as performance pay may have (see (Gimpelson et al. 2009) for a detailed discussion of salaries in the public sector in Russia).

Performance management as a floating signifier

> *Significant divergences between talk, decisions and actions are often the norm in organizational life (Brunsson 1989, cited in (Pollitt 2001: 935))*

What, then, can we gain by looking at "performance management" as a "floating signifier"? First, it lets us understand that there is no essential link between performance indicators *per se* and any particular use to which they can be put. NPM doctrine suggests that performance management should be associated with managerial devolution, but in Russia performance indicators were used as a tool of centralization and administrative regulation.

Second, similarity of form does not mean similarity of content. Performance management in Britain used indicators, league tables, rankings, rewards and targets. In Russia, all the same elements were used but acting agents attached somewhat different meaning to them. Despite evident convergence on the level of "ideas" and "policy rhetoric" there was little convergence in actual practice. This is only one of many possible levels of convergence:

> *… convergence might be on policy rhetoric – or what is sometimes called 'talk' or discourse (Pollitt 2001). Phrases, buzzwords and clichés can take a life of their own– often one that is somewhat content-less and divorced from theoretical underpinnings. Some might be phrases that all can agree to support, without necessarily agreeing on what they mean.(Goldfinch and Wallis 2010: 1101)*

Three, the form may itself serve a function. Using performance management as a brand allows policy-makers to create a certain impression of innovativeness and focus discussion on problems that are common for developed countries. Thus, an illusion of modernity and civilization is created: a veil of progress that covers up the rotten core with home-grown problems and illnesses.

Four, the form allows for intricate interplay of power relations to take place. Allegiance to the policy of performance management may be used as an indicator of trustworthiness of a civil servant. If one is able to propagate the fiction of performance management, then one is a good candidate for top positions and can be trusted other secrets of power. Ideologies may be used as indicators of loyalty.

Performance indicators as a rhetorical tool

> *…a great deal of* talk *can develop around concepts such as the New Public Management (NPM) without that necessarily signifying an equivalent amount of action in the same direction (Pollitt 2001: 934)*

Another type of usage of performance indicators was observed during fieldwork. At a public hearing, where a new socio-economic strategy was discussed, members of the public complained that they knew little about effectiveness of budget spending. It was implied in the question that the transparency of budget spending is insufficient and that it is hard for the members of the public to decide how regional administration was performing. In reply to this reproach one of the representatives of the Ministry of economic development said that there is abundant information on the results achieved by the government. He pointed out that over 300 indicators were published annually and that they provided sufficient information on all relevant aspects of public sector effectiveness. When one of civil activists tried to reply that the format of indicators was inaccessible to the majority of the population, the civil servant answered that they are accessible to anyone who has basic reading skills. This case becomes interesting when it is noted that this civil servant only a day before was interviewed by the researcher and was one of the fiercest critics of the system of performance measurement. He held that performance indicators were an empty red-tape and did not reflect real government performance. He said that they were used only as a "club to hit governors on the head" (meaning that among 300 indicators there were always some underperformers which made it easier for the president to bargain with governors). But a day later this same person publicly argued that these indicators provided abundant information and could be used to monitor regional government's performance.

This case can be interpreted as an example of one of symbolic functions of performance indicators. Similarly to politicians, who resort to "objective measures" to justify arbitrary political choices, civil servants resort to "detailed performance measures" when they want to deflect criticism, avoid blame and safe their face. Face saving is one of many ritual elements in social interactions (Goffman 2006). In the case under consideration the person who had to perform the ritual was not speaking for himself as an individual, but was speaking on behalf of his organisation and the government in general. Therefore, he had to make a "face saving" move to protect his group from accusations of insufficient transparency, even though he himself did not believe what he was saying to be true. This "face saving" function of indicators could be an important one and could partially explain why indicators with little direct managerial value are kept and used.

The role of performance management as a rhetorical device should not be underestimated. As other policies, it has the power to test loyalty of civil servants and experts. It may be an important device in bonding together the community of experts, even if all it generates is talk rather than actual practical reforms:

> *Talk has a life of its own ... It is a means of winning legitimacy and support, but not only that. It creates an agenda of issues and a specialist vocabulary for describing them. It facilitates the formation of 'in groups' (who share the agenda and vocabulary) and 'out groups' (who don't). Within the relevant 'community of discourse', deployment of the current concepts and 'buzzwords' confirm the speaker as 'knowing' and 'on message'. By the same token, discussants who raise other issues or use an older vocabulary may be disempowered – regarded by those 'in the know' as embarrassingly 'old-fashioned', underinformed or even subversive (Pollitt 2001: 939).*

Performance data as a façade of kleptocratic bureaucracy

A different interpretation may be given to the seeming prevalence of *prudent bureaucrat* behaviour. It may be linked to the dominance of rent-seeking among motives that bring people into the civil service. Studies of the civil service in Russia show that corruption and abuse of office are widespread. Orlova (2008) observes that despite recent harsh political rhetoric and a few high-profile scandals 'corruption still flourishes. In public administration … [a] semi-feudal

chain of command that profits from bribery has been established'. From this perspective, the strive for greater security and convenience that characterizes *conservers* in Downs' typology may be interpreted as rent-seekers' desire to minimize the burden of official duties and free up more time for abuse of office. The civil service as a whole and performance reports in particular may then be considered "a façade" concealing the true nature of bureaucratic activities, that is, extraction of bribes. It is natural to expect such corrupt bureaucrats (kleptocrats) to try to remain as inconspicuous as possible and maximize their long-term revenue by staying in office as long as possible and doing as little as possible. The difference between climbers and conservers in kleptocratic terms then would be in their strategies of maximizing rent. *Climbers* prefer to advance in the hierarchy to reach the position where they can get much larger bribes. *Conservers*, in contrast, prefer taking smaller bribes over a longer period of time. For arguments for and against the interpretation of Russia as a kleptocratic state see (Dawisha 2011).

Performance data as a by-product of managerial work

Longenecker and Ludwig (1990) make an important distinction between the use of performance data as an end in itself (this is the way the data are used by Human Resource departments in private organisations or by performance management units in governments) and as means to obtain other ends (this is the way performance data are used by managers or public servants who generate it). In the first case, the data are valuable because they provides the material for decision-making by the principal; in the latter case performance data are used to achieve some goals other than higher performance in its narrow sense. Managers may use performance data to formally justify rewards to selected subordinates or to provide official backing for politically motivated decisions. Performance figures are then just by-products of a wider activity. Whatever the reason, the result is that accuracy of performance data is undermined.

Moreover, from a public sector manager's point of view the very fact that performance data are ambiguous may be an advantage. Ambiguous figures may be used to rationalise discretionary decisions. As one consultant observed:

> *Among three hundred indicators the governor can always find those that can be used to either reward a minister or to dismiss him.*

In addition, this usage of performance indicators is analogous to the use of tax legislation by "blackmail states" (Darden 2008). The idea of a blackmail state is that it purposefully passes contradictory tax legislation. This makes compliance nearly impossible and forces agents to evade taxes. Economic actors are then blackmailed to ensure their political support. Similarly, the sheer number of performance indicators opens opportunities for blackmailing.

Conclusion

Implementation as symptom

The implementation of performance indicators was only one new policy in the constant flow of new government initiatives. It has, however, inevitably left its mark on the organisational culture and procedures within organisation. Successes, failures and problems that emerged during the implementation of this policy are indicative of the quality of public sector management in general. Effectiveness and vigour with which this major policy has been implemented (or lack of such vigour) illustrates the general commitment of the reformers to steer in a particular direction. By observing the implementation process of a policy one can make inferences about the context in which the policy is implemented.

In the Russian case, for example, the fact that performance management has not survived the competition with conflicting reforms may mean that the context was not ready for reforms focusing on *sigma*-values. Other reform initiatives (such as administrative regulation, standardization and bureaucratic transparency) fitted the administrative environment better and thus were able to take deeper roots.

Performance measurement in Russia was implemented in isolation from other NPM initiatives. No managerial devolution accompanied the implementation of performance indicators. As a result, public sector managers were turned into mere administrators and their successful performance (or lack of it) depended more on external factors than on their own efforts. In this context indicators were "labour-optimized": in many cases data were manipulated to comply with desirable trends in either "prudent" or "reckless" way. The sheer abundance of information (over 300 indicators) interfered with effective use of the data. Useful information was bogged down in the morass of cumbersome performance indicators.

Despite the fact that performance indicators were perceived as delivering few managerial benefits they continued to be used by civil servants. This may indicate that there were other benefits such as political and rhetorical benefits that supported the continual use of the system.

Despite seeming convergence with global NPM agenda the implementation of performance indicators in Russia was different in content despite being similar in form. Indicators were introduced as an instrument of bureaucratic accountability with no provisions for devolution of managerial decision-making. This is an example of a more general problem. Convergence of rhetoric is not necessarily linked to convergence in actual practice:

> *Convergence on rhetoric can act to obscure a variety of design and policy decisions, and new language and ideas are filtered, resisted and adapted through the differing institutions of different countries to produce quite different decisions and outcomes. Implementation of apparently similar decisions and instruments might be associated with huge divergence in actual practice (Goldfinch and Wallis 2010: 1104)*

CHAPTER 9. CONCLUSION

Those who study performance management from the academic perspective should try to avoid judging it too strictly. As a technocratic tool, performance management/measurement has its limitations and cannot become a substitute for good governance. Moynihan (2008) mentions two of these unavoidable limitations:

> *… performance information systems produce mounds of information that no one particularly cares about and that collectively is beyond the cognitive abilities of any individual to process. Another criticism is that politics makes performance information irrelevant. Strong political preferences make performance information unnecessary. Relative to partisan goals, ideological biases, stakeholder pressure, and constituent needs, performance data are not especially influential. In addition, performance information does not help elected officials by making political decisions simpler indeed, it is an additional layer of information to incorporate. (Moynihan 2008: 11)*

For the Russian case it should particularly be kept in mind that PM is not a panacea and that weak implementation is not a uniquely Russian problem. Other countries have faced similar problems. The Russian case should not be thought of as somehow unique, but, on the contrary, may be regarded as an example of general problems of implementing PM. In countries with similar social-economic and political climate PM is likely to adopt similar shapes. Findings of this study would be particularly useful for researchers and practitioners from developing countries as they are likely to face similar problems as Russia.

Kelly observed that "reforms leave footprints" and even if they are not successful, they increase our understanding of the public sector. She also observes that Performance management movement has shown that the task of linking budgeting with performance is unrealistic, though noble:

> *In my home field of public budgeting, we have learned that fleeting reform movements leave "footprints" on the profession long after another reform movement has made the past one a faint memory. <…> And alas, the current research on our latest reform, performance budgeting, indicates that performance does not have much effect on budgeting. But we are learning something about how to make the budget document an accountability document by dealing explicitly with performance issues. <…> By our trying and our failing, we develop as a field, if not exactly as a science. (Kelly 2002: 329)*

Lessons learned from the Russian experience may be relevant for implementers in other developing countries. In the Russia performance measures have been implemented within a rich network of informal incentives that made data manipulation likely. A particular type of data manipulation ("prudent bureaucrat" manipulation) was registered through both interviews and statistical analysis.

In the Russian case the implementation of performance measurement instruments was not accompanied by fiscal devolution. As a result, performance reporting received relatively little attention from managers and was "labour-optimised". Managers were turned into mere administrators. Available evidence suggests that data for a number of indicators were made up to conform to a "prudent" trend, presumably to avoid the trouble of collecting actual data and avoid unnecessary attention from supervising bodies. The excessive number of indicators (over 300) made the system cumbersome for civil servants. Useful data were sometimes lost in the sheer volume of statistics.

The following major problems were identified: lack of interagency cooperation, lack of feedback during the process of implementation, inability of agencies to influence performance trends in the short run coupled with short-term based rewards and sanctions. At the same time it was found that respondents noted some managerial benefits of the existing system: cross-regional benchmarking was instrumental in identifying areas of governmental failure. Indicators were also used to signal demand for funding to a service.

What have we learned?

Four research questions were asked in the beginning of this thesis:

1. How do formal performance indicators function in the context of the Russian regional government?
2. What managerial benefits do civil servants see in the system?
3. Is the system of performance management susceptible to deliberate data manipulation?
4. If yes, then what type of data manipulation is most common?

By answering these four questions we can learn more about performance management as a public sector reform, about the process of its implementation, about bureaucracy in general and about Russian bureaucracy in particular.

What, then, can we answer to these questions now?

How do formal performance indicators function in the context of the Russian regional government?

It has been found that performance indicators perform a number of functions in the practice of regional civil servants, these may be grouped in four groups: managerial, political, rhetorical and ideological. Chapter 5 provided qualitative evidence for these different types of usage. First, indicators are used for managerial purposes: it has been found that they are used for benchmarking between regions and also for identifying areas of failure. League tables are instrumental in this respect. It is important to note that performance measurement system served some symbolic purposes – such as the purpose of creating an impression of rationality over government's decisions. Overall, the system of formal indicators became entangled in the pre-existing network of tacit rules. As a result, incentives were created that were not explicitly envisaged by those who designed the system. These informal rules resulted in significant pressure on civil servants to demonstrate improvement of some performance indicators.

The only official reward for improving indicators was the mechanism of ranking regions and distributing grants. This system was not sufficiently predictable and did not provide large enough incentives to make governors pay sustained attention to improving performance indicators. In addition to low formal incentives the number of indicators was too high (over 300) to make any concentrated effort possible. As a result, the process of performance reporting was "labour-optimized": instead of collecting data and monitoring actual performance some indicators were simply made up to comply with a hypothetical neutral trend and avoid attention of supervising authorities. Overall, it may be confidently said that the current system did not

produce a change in public sector mentality and “bobbed along” on the surface of government practices.

The system also generated benefits that may be called ideological. Among 325 indicators one could always find a few indicators that improved and use them as a selective example. Another ideological function was the testing of loyalty of civil servants. It was observed that the same civil servant in private conversations with the interviewer criticized the system of performance management as useless and wasteful, but when this same person took part in a public hearing she spoke confidently in defence of the system against accusations made by civic activists and citizens. Civil servants do not believe that the system is functioning well, but defend it against criticism of others. So the system provides a tool of selection from among civil servants on the criteria of being able to hypocritically defend it in public. This is an example where selection against honesty takes place.

What managerial benefits do civil servants see in the system?

It was found that indicators do not perform many functions that are traditionally assigned to them by public sector management literature. The table below shows that for six out of eight managerial functions that are traditionally mentioned as reasons for implementing performance indicators only marginal benefit could be identified.

Table 25. Perceived managerial benefits of performance indicators in the Russian case. The list of functions adapted from Behn (2003).

Managerial function	Meaning	Perceived managerial benefit	Reason
Evaluate	How well is my public agency performing?	Marginal	Agency's contribution to the dynamics of indicators is small: external effects, long-termism, lack of cooperation
Control	How can I ensure that my subordinates are doing the right thing?	Marginal	
Budget	On what programmes, people, or projects should my agency spend the public's money?	Marginal	Determining role of federal decisions. Compulsory funding of federal mandates. Little local discretion.
Motivate	How can I motivate line staff, middle managers, non-profit and for-profit collaborators, stakeholders, and citizens to do the things necessary to improve performance?	Marginal	Alienation. Indicators are handed down from the top. Lack of involvement, communication, feedback and local initiative.
Promote	How can I convince political superiors, legislators, stakeholders, journalists, and citizens that my agency is doing a good job?	Marginal managerial benefit. Substantial symbolic benefits.	This may be regarded as a political (symbolic) benefit. Indicators create the veil of technocratic rationality.
Celebrate	What accomplishments are worthy of the important organizational ritual of celebrating success?	Marginal	Agency's contribution to the dynamics of indicators is small: external effects, long-termism, lack of cooperation
Learn	Why is what working or not working?	Some benefit identified	Cross-regional comparison may be useful for problem diagnosis, but lack of continuity in performance management systems undermines potential benefits for learning over time.
Improve	What exactly should who do differently to improve performance?	Marginal	Lack of interagency cooperation. Fragmented and rigid responsibility.

Respondents emphasized informal symbolic functions of performance indicators over their official functions. It was, for example, noted that performance indicators play an important role in the competition between governors both through formal and informal channels. Another symbolic benefit of the system was its ability to generate the impression of novelty and innovation demanded by politicians.

At the same time it was observed that indicators served as a tool of bureaucratic control in the administrative hierarchy. This is something that the system was not officially intended to do. Lower civil servants were monitored on indicators that officially were implemented as mere *intelligence* measures and were not to be used as targets or for ranking.

Is the system of performance management susceptible to deliberate data manipulation?

According to respondents, cases of data manipulation were common in the system. For some indicators objective data could not be collected because of resource and capacity constraints, other indicators were deliberately manipulated to demonstrate favourable performance. In addition, there was no systematic audit, so cheating was not penalized.

If yes, then what type of data manipulation is most common?

According to quantitative findings the scope of manipulation differed depending on the agency responsible for data collection. Two types of data manipulation strategies were identified: "prudent bureaucrat" and "reckless bureaucrat" data manipulation. The first consisted in smoothing out annual fluctuations in data and reporting a neutral non-suspicious trend. The second consisted in consistent inflation of performance data. Prudent data manipulation was dominant among indicators that were not included in league tables. Most obvious traces of this type of manipulation were observed in the group of self-reported indicators collected by regional administrations.

In both cases it is difficult to separate effects of cheating from honest figures, but an attempt was made to estimate the "natural" variation of performance indicators. It has been shown through a survey of municipal civil servants that self-reported indicators do not seem to

be different in their natural variation from the rest of the indicators and, thus, a large proportion of the abnormal dynamics may be attributed to measurement bias or cheating. Respondents of the survey perceived indicators with high rate of clustering in the vicinity of zero as less trustworthy. A statistically significant correlation was demonstrated to exist between the perceived trustworthiness and index of clustering calculated as the share of observations in the vicinity of zero (range from -0,3% to 0,3%). The very attempt to demonstrate inconspicuous behaviour by many reporting agents created suspicious trends in overall data. A possible interpretations is that respondents knew from experience that it was problematic to obtain good quality data for these indicators. Because of this, they were less inclined to trust the data for these indicators reported by others.

The evidence for reckless manipulation appears inconclusive. On the one hand, it has been shown that in some years indicators tied to grants reported by Rosstat demonstrated significantly better performance. On the other hand, indicators reported by Federal ministries do not seem to comply to the hypothesized trend. It appears that grants did not create sufficient and consistent incentives for reckless manipulation. The largest irregularity was shown to exist in year 2008 – the second year of the working of the system. Perhaps this was due to the fact that reporting agents were for the first time able to compare themselves with others and against national averages. This may have resulted in active manipulation to catch up with the rest of the regions. In subsequent years there does not seem to be a sustained effect of "reckless manipulation". It appears that incentives created by league tables did not lead to mass gaming.

Theoretical implications

Misrepresentation is a problem that any performance measurement regime should acknowledge. It is important to understand incentives that lead to data manipulation. It is often implied or explicitly theorized that blame avoidance in public sector performance reporting stems from the fear of scandals and close attention of the public (Davies 2004; Hood 2011). Our findings show that, even in a performance measurement regime insulated from both the public

and the legislature, blame avoidance seems to have a significant effect on performance measures. Bureaucratic environment itself sustains the blame avoidance culture. Upward accountability creates stimuli for keeping a low profile.

Another important finding is related to data manipulation practice among indicators operating in *intelligence* mode. Hood (2007b: 96) observed that the practice of intelligence regimes was underresearched and that the use of *intelligence* regimes "is unpredictable by those whose performance is recorded". We can now add some clarity to this question. In the absence of pressure of targets or rankings, agents resort to a cost-minimisation strategy (what we termed here as a "prudent" data manipulation strategy). Moreover, in the absence of formal audit this strategy works for cost-minimisation of both principals and agents. Prudently manipulated data are inconspicuous to a supervising authority. It requires some meta-analysis as one offered in this thesis to render these distortions visible. Hood (2007b: 96) Hood (2007b: 96)

If no provisions are made to check distorting behaviour of reporting agents, technical usefulness of a performance measurement regime as a nation-wide tool of benchmarking and administrative decision-making may be undermined. It may well be, however, that accuracy of performance data is not the real aim of such a regime. A variety of motives can motivate central governments to implement a "meaningless system" overloaded with data. One such motive may be the normalising effect of performance reporting. By putting administrative men in a panopticon the supervisor can enforce conformity and exert power. The fact that the resulting data become distorted may not be regarded as an issue, as it is not the data that are valuable for the supervisor, but the process of reporting itself, as one respondent observed:

> *Among three hundred indicators a governor can always find those that can be used to justify either rewarding a minister or dismissing him.*

A prudent bureaucrat in such a system finds ways to minimize his exposure to blame, one such way is data manipulation.

Contribution of the study

The study contributes to different streams of literature. The first stream is concerned with implementation studies. A number of deficiencies have been identified that hindered the implementation and subsequent evolution of a performance measurement system. The Russian case is a practical demonstration of theoretical prescriptions. Lack of feedback, strict top-down implementation, insufficient attention to building local capacity were all demonstrated to be present in the Russian case. The study, thus, supports overall prescriptions generated by implementation studies.

The second stream to which the study contributes consists in studies devoted to the study of performance management as a managerial practice. Performance management can be both a substantial reform and a decorative element. This is, probably, true of any reform. In the Russian case performance management has not yet been transformed into a substantive reform effort. It has been existing nominally but has not received attention and resources needed to transform it into a tangible element of public reform arena. It has been bobbing along the surface and had only a passing and superficial effect on daily work of civil servants. Everyone knows that performance indicators exist, but few know what they are for. Nonetheless, resources (both money and time) were spent to design and roll out the system. No systematic effort has been made to assess the costs of implementing and keeping the system. The costs of the system may exceed its benefits.

The third stream is concerned with particularities of the development of the public sector in Russia. The contribution here is in demonstrating how performance management initiative was implemented in the environment where the "capital base" of public sector values is yet to be built. It may be that in the Russian case there was a mismatch between values promoted by the reform and values for which there existed substantial demand. In theory, NPM policies are focused on promoting *sigma* values of economy, parsimony and effectiveness, but in the Russian case there seemed to be more demand for *lambda*-values of order and dependability and, more

recently, the pressure to promote *theta* values of transparency. In this case performance management was implemented in the kind of environment that has not yet become mature enough for it. In the Russian case performance management was implemented in a non-NPM mode, it was not accompanied by managerial devolution and other traditional elements of NPM. The capital base of a smoothly working public sector bureaucracy is yet to be built in Russia. The shallowness of performance management reform itself acts as an indicator of immaturity of the public sector. Compared to western practice, the implementation of "performance management" in the Russian case may be interpreted as a "[certain type of] convergence, where similar external labels or titles are used for reforms, although the substantive content of these reforms may vary considerably" (Pollitt 2001: 944).

Fourth, this study contributes to the literature on public sector reform in developing countries. Apparently, there may be PM without NPM and NPM without PM. The link between the two is not cut in stone. This may be relevant for non-western contexts. In the Russian case, conflicting ends were pursued by different elements of public sector reform package. On the one hand benchmarking was introduced to stimulate competition between regions, and on the other hand, a plethora of very detailed performance measures was introduced to monitor inputs, processes and outputs. Such detailed control contradicted the idea of devolution of managerial discretion and created a cumbersome system that generated a lot of low quality data. Perhaps, this is a necessary step precluding any "real" implementation of "serious" performance management and is necessary in the process of maturing of the public sector. Moreover, in Russia there exists a conflict between federalism and comprehensive performance management system at the national level. A rigid top-down performance management system may violate the idea of federalism.

Fifth, the study contributes to research on how bureaucracies operate. The "prudent bureaucrat" behaviour is theoretically linked to the theory of bounded rationality of Herbert Simon and the theory of normalization of Michel Foucault. The tendency to try to stay "under

the radar" leads to a peculiar trend in performance data. If data are not properly audited then there is a danger of manipulation. Bureaucrats tend to report minor variations of performance to avoid both sharp increases and sharp drops in performance because these can attract unwelcome attention from supervisors.

Finally, the study contributes to research on the effects of monetary incentives on performance. It has been shown that indicators linked to the league table and performance-based grants behave differently from indicators not tied to grants. Theoretically, league tables can generate psychological incentives and even if the monetary prize is relatively small the psychological effect may be considerable. Everybody wants to have a higher place in a league table. So, even without explicit targets, league tables can generate upward pressure on indicators. In the Russian case, the monetary stimulus of federal grants was relatively small. The psychological effect, on the contrary could be considerable. On the other hand, respondents noted that the methodology of distributing grants was too unpredictable and it was not possible to strategically plan for improved performance in the league table. This indicates that the psychological effect was underutilized. If accompanied by clear rules and consistent methodology, it could be a more potent tool. In the Russian case league tables every year came as a surprise for civil servants and thus there were no incentives to implement development plans to achieve higher places in ranking. The list of top 10 regions persisted almost unchanged between 2009 and 2011, leaving little chance for the majority of regions to participate in the race for federal grants. Perhaps a different type of league tables with more differentiated rewards would perform better. Agencies could be given star ratings to engage them into the game of substantial improvement. This opens up opportunities for further research into the effect of different types of rankings.

Qualitative evidence shows that some indicators generated informal pressure for improvement through the chain of principal-agent relations. Those at the bottom were assessed by their superiors on the basis of these indicators and were pressured by informal incentives to

demonstrate sustained improvement (or, at least, to avoid reporting minor negative growth). In the climate of the public sector with its strong blame avoidance bias this potentially could create conditions for negative selection against honesty. This may explain the dominance of "prudent manipulation" with significant drop in small negative values of growth indices. Indicators were "massaged" to avoid demonstrating small negative growth. This did not lead to significant reckless manipulations, however. This may indicate that, in Downs' terminology, *conservers* constitute the majority among civil servants or that the reward system does not provide enough perks for sustainable improvement. Perhaps, this indicates that most Russian bureaucrats prefer security over ambition.

The results of the validation survey suggest that indicators with pronounced "prudent bureaucrat" behaviour (those for which growth rates were clustered in the vicinity of zero) are perceived as less trustworthy. A possible explanation is that civil servants know from experience that these stagnating indicators are often "drawn" and "massaged" to reach the zero-threshold and project this knowledge on indicators collected by other municipalities and regions.

Policy implications

Policy recommendations are, in general, in line with the implementation studies of performance management regimes (Carter 1991; Likierman 1993; Meekings 1995; Norman 2002; Ryan 2004).

Policy-makers may find it useful to apply the method of detecting data manipulation in assessing the quality of performance measurement regimes. Traces of "prudent" data manipulation may indicate that the system of rewards does not create incentives for sustainable improvement of performance. On the other hand, if indicators tied to rewards demonstrate outstanding improvement it should be taken cautiously because it may indicate that "reckless manipulation" is taking place.

Trust in performance data may be improved by allocating greater resources to data collection and audit. It should be taken as rule of thumb that it is better to have a few robust indicators than a plethora of low quality measures.

Informal rules should be taken into consideration when designing systems of performance management. Indicators that are formally implemented in *intelligence* mode may spontaneously generate unintended informal incentives at different levels of the bureaucratic hierarchy. The fact that there is no formal target for an indicators should not be taken to mean that this indicator has no distortionary effect on managerial behaviour.

Efforts should be made to formalize the procedure of rewarding and sanctioning, but it should be remembered that informal incentives can never be fully removed. The study has shown that frequent changes in methodology of ranking are detrimental to the functioning of performance management regime. People need time to get used to the rules of the game. If these rules change too frequently people stop taking the system seriously and shift to mechanical compliance.

The Russian case shows that, if performance measurement reform is not linked to a wider reform process in the public sector, it cannot produce deep change. In the Russian case other elements of NPM (such as performance pay, freedom to manage and devolution) were not consistently implemented and thus performance indicators became an isolated and unsupported practice. This gradually led to its marginalization and decline. Performance reporting became a hollow compliance exercise separated from day-to-day practice of civil servants.

It was found that performance management as a separate reform effort was implemented in large part for its reputation as a global public administration trend. Respondents remarked that the image of innovativeness that came with performance management was important. But the reform could not achieve its declared aims in isolation from other reform initiatives. In the Russian case such consistency was not reached (Verheijen and Dobrolyubova 2007). As a result,

performance measurement remained an artificial element that "bobbed on the surface" of public sector practice.

These findings may be generalized to other contexts, particularly non-Western. It is likely to be relevant for other countries with similar political regimes. Russia has been characterized by North et al. (2009: 31) as a limited access order society. Most counties in the modern world are limited access order societies. The relevance of the findings for Western contexts may be low, because of radically different relationships between the state and the society and because of established tradition of public sector integrity. The findings are probably most relevant for societies where trappings of democracy exit. In such societies there is likely to be a demand for novel western managerial ideas. If there is a need to demonstrate that the public sector is innovative and modern, performance management is one of many policies that may be used to create an impression of innovation and modernization. It may be useful as a symbol if not as a real managerial practice.

APPENDICES

Appendix 1. Interview guide

English version

In brackets – original numbering in Pollitt (2006)

A. How is ‘‘performance’’ conceived of and measured in the sample of agencies studied?

1. (3). If I say to you ‘‘the performance of X [agency name here],’’ what are the first aspects that spring to mind?
2. (4). What are the most important measures or indicators, from the point of view of agency staff?

B. Have there been any significant recent changes in the way in which performance is thought about here?

1. (6). Have these changes been reflected in the measures and indicators?
2. (7). Who determines what the current set of performance measures should be? (Where appropriate, pursue this with supplementary questions aimed at establishing how far the agency effectively sets its own indicators or, alternatively, how far indicators are imposed on it from outside.)

C. How Far, by Whom, and for What Are the Performance Data Used?

1. (8). Are performance data much discussed within the agency?
2. (9). Who makes the most use of them inside the agency?
3. (10). Who (if anyone) makes the most use of them outside the agency? (Supplementary: Can you give an example of their use?)

D. To What Extent Are Performance Measurement Systems Integrated with Financial Management Systems? Where This Integration Is Low, What Are the Reasons for That?

1. (13). How far do you think the performance data reflects the really important aspects of the agency’s ‘real’ performance?
2. (14). In general, how closely is the performance measurement system linked in with the other main management systems in the agency?
3. (15). What kinds of links are there with financial systems? (If weak, then why?)
4. (16). What kinds of links are there with planning systems? (If weak, then why?)
5. (17). What kinds of links are there with human resource management systems? (If weak, then why?)

E. External/Contextual Influences on Performance

1. (24). What are the main external influences on how well the agency performs?(Supplementary: Through what process do these influences operate?)

Russian version

В скобках указана нумерация в оригинальном интервью-гиде Поллитт (2006)

A. Как «результативность» понимается и измеряется в исследуемых организациях?

1. (3). Если я скажу вам «результативность X [название ведомства]», какие характеристики вам приходят на ум прежде всего?
2. (4). Какие индикаторы или показатели являются наиболее важными с точки зрения сотрудников организации?

B. В последнее время произошли ли какие-либо серьёзные изменения в том, как результативность понимается в организации?

1. (6). Были ли эти изменения отражены в используемых индикаторах и показателях?
2. (7). Кто определяет, каким должен быть используемый набор показателей результативности? (В случае необходимости, следует задать дополнительные вопросы, чтобы установить, в какой мере агентство устанавливает свои собственные индикаторы или, наоборот, насколько показатели, устанавливаются для агентства со стороны.)

C. Кем и для чего используются данные о результативности?

1. (8). Обсуждаются ли данные о результативности внутри организации?
2. (9). Кто в наибольшей степени использует данные о результативности внутри организации?
3. (10). Кто (если кто-либо) в основном использует данные о результативности за пределами агентства? (Дополнительно: Можете ли вы привести пример их использования?)
4. Обсуждаются ли показатели результативности в местной прессе/жителями?

D. В какой степени показатели результативности увязаны с системой бюджетирования? Где это увязаны слабо, каковы причины этого?

1. (13). Как вы думаете, насколько данные о результативности отражают действительно важные стороны "реальной" работы организации?
2. (14). В целом, насколько тесно система оценки результативности увязана с другими системами управления в организации?
3. (15). Какова увязка с финансовой системой? (Если слабая, то почему?)
4. (16). Какова увязка с системой планирования? (Если слабая, то почему?)
5. (17). Какова увязка с системой управления кадрами? (Если слабая, то почему?)

E. Внешние факторы, влияющие на результативность

1. (24). Каковы основные внешние воздействия, влияющие на работу организации? (Дополнительно: каким образом действуют эти внешние факторы?)

Appendix 2. Validation questionnaire

English

Respondents were given these 5 questions to answer:

1. In your opinion, how significant is the influence of municipal governments over a given indicator?

0% 25% 50% 75% 100%

0% - the dynamics of this indicator is fully dependent on other levels of government;
100% - the dynamics of this indicator is fully dependent on the municipal government.

2. In your opinion, if the government (local and regional) does not intervene, what would be the dynamics of this indicator in during the next 3-5 years?

- will rise sharply,
- will rise slowly,
- will remain stable,
- will fluctuate moderately,
- will fluctuate widely,
- will decrease slowly,
- will decrease sharply,
- do not know.

3. In your opinion, if the government had abundant resources and could allocate them to the area measured by this indicator, how easy would it be to achieve growth of 5% in a year?

- difficult;
- extremely difficult;
- possible without significant problems;
- can be easily achieved;
- do not know.

4. In your opinion, is your municipality able to collect objective and accurate data on the value of this indicator? Do you agree with the following statement – "Our municipality can collect accurate and reliable data to measure values of this indicator".

- Strongly agree
- Agree
- Disagree
- Strongly disagree
- Do not know

5. In your opinion, how accurate and objective is the data on values of this indicator, reported by other municipalities in your region? Do you agree with the following statement – "Other municipalities of my region provide reliable and accurate data on values of this indicator".

- Strongly agree
- Agree
- Disagree
- Strongly disagree

Russian

1. Оцените степень влияния органов местного самоуправления на изменения значений показателя.

   0% 25% 50% 75% 100%

   0% - динамика показателя полностью определяется другими органами государственного управления;
   100% - динамика показателя находится под полным контролем нашего ОМСУ

2. Если органы государственного управления и органы местного самоуправления не будут принимать никаких усилий по управлению соответствующей сферой, какую динамику, по-Вашему, будет демонстрировать показатель в течение следующих трёх-пяти лет?

- будет стремительно расти;
- будет медленно расти;
- будет стремительно снижаться;
- будет медленно снижаться;
- будет значительно колебаться от года к году
- будет незначительно колебаться от года к году
- затрудняюсь ответить.

3. Предположим, что в распоряжении органов государственного управления имеются очень значительные свободные ресурсы (финансовые, кадровые и т.д.) для развития данной сферы. Насколько трудно в такой ситуации добиться улучшения значения показателя на 5% за год? Закончите фразу: "Добиться улучшения значения показателя на 5% за год…

- невозможно;
- крайне сложно;
- можно;
- легко.
- затрудняюсь ответить.

4. Согласны ли вы со следующим утверждением: «Для оценки значения этого показателя ОМСУ могут собрать достоверную информацию; исходная информация для расчёта или информация о значениях показателя, предоставляемая иными органами власти, достоверна, получена из надёжных источников».

- полностью согласен
- скорее согласен
- скорее не согласен
- совершенно не согласен
- затрудняюсь ответить.

5. Согласны ли вы со следующим утверждением: «Данные о значениях этого показателя, предоставляемые в Докладах глав других муниципальных образований, расположенных на территории моего субъекта РФ, достоверны. Этим данным можно доверять»?

- полностью согласен
- скорее согласен
- скорее не согласен
- совершенно не согласен
- затрудняюсь ответить.

Appendix 3. List of reporting agencies

Table 26. The sources of performance data in the national dataset.

Reporting agency	Number of indicators (as of 2011)
Federal Bureau of Statistics (Rosstat)	75
Regional Administrations	72
Federal Ministry of Health	46
Federal Ministry of Education	36
Federal Treasury	30
Federal Ministry of Regional Development	13
Federal Security Service (FSO)	8
Federal Ministry of Sport	8
Indicators that were re-assigned during the period	8
Federal Service for Control in Education (Rosobrnadzor)	7
Federal Ministry of Finance	6
Federal Ministry of Internal Affairs	5
Federal Ministry of Natural Resources	4
Federal Ministry of Economic Development	3
Federal Ministry of Agriculture	1
Federal Medical Insurance Fund	1
Federal Ministry of Culture	1
Federal Service for Control over Consumers' Goods (Rospotrebnadzor)	1
Total number of indicators	325

Appendix 4. Descriptive statistics

Average, median and mathematical expectancy for distributions of growth rates for four groups: self-reported indicators, Rosstat, Federal ministries and FSO. Values within the range from -1 to 1 are reported to avoid the bias of outliers and errors in the database.

Mean growth rate

Mean	Rosstat	**Self-reported**	Fed. ministries	FSO
2008	0,018	**-0,008**	-0,018	0,024
2009	0,006	**-0,040**	-0,016	0,062
2010	0,016	**-0,054**	-0,035	0,039
2011	0,017	**0,013**	0,001	0,003
All years	0,017	**-0,008**	-0,018	0,024

Median growth rate

Median	Rosstat	**Self-reported**	Fed. ministries	FSO
2008	0,014	**-0,010**	-0,021	0,007
2009	-0,002	**-0,038**	-0,030	0,042
2010	0,004	**-0,046**	-0,038	0,033
2011	0,016	**0,011**	0,000	-0,007
All years	0,013	**-0,010**	-0,021	0,007

Mathematical expectancy of growth rate

Math. Exp.	Rosstat	**Self-reported**	Fed. ministries	FSO
2008	-0,109	**-0,019**	0,557	0,170
2009	-0,001	**-0,053**	-0,025	0,057
2010	0,007	**-0,064**	-0,042	0,034
2011	0,011	**0,004**	-0,005	-0,002
All years	0,010	**-0,018**	-0,025	0,019

Appendix 5. Share of observations clustered in the vicinity of zero

The tables below report the share of observations clustered in narrow intervals around zero. Growth rates of self-reported indicators tend to be most convergent to zero (have greater kurtosis values).

2008

Interval	Rosstat	**Self-reported**	Fed. ministries	FSO
-0,05 - 0,05	0,28	**0,39**	0,32	0,43
-0,03 - 0,03	0,20	**0,34**	0,24	0,36
-0,02 - 0,02	0,16	**0,31**	0,20	0,32
-0,01 - 0,01	0,12	**0,28**	0,14	0,28
0 (kurtosis)	0,07	**0,25**	0,08	0,17

2009

Interval	Rosstat	**Self-reported**	Fed. ministries	FSO
-0,05 - 0,05	0,35	**0,38**	0,36	0,28
-0,03 - 0,03	0,26	**0,33**	0,27	0,18
-0,02 - 0,02	0,21	**0,30**	0,22	0,13
-0,01 - 0,01	0,16	**0,27**	0,15	0,09
0 (kurtosis)	0,09	**0,23**	0,08	0,03

2010

Interval	Rosstat	**Self-reported**	Fed. ministries	FSO
-0,05 - 0,05	0,39	**0,36**	0,38	0,34
-0,03 - 0,03	0,28	**0,31**	0,27	0,23
-0,02 - 0,02	0,21	**0,28**	0,21	0,16
-0,01 - 0,01	0,15	**0,24**	0,14	0,09
0 (kurtosis)	0,08	**0,20**	0,06	0,03

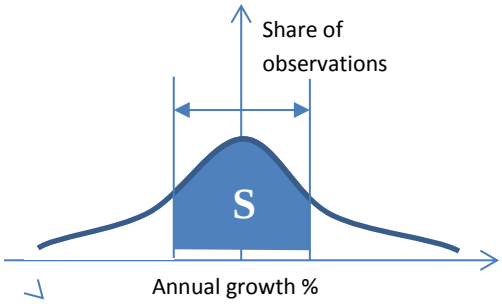
2011

Interval	Rosstat	**Self-reported**	Fed. ministries	FSO
-0,05 - 0,05	0,38	**0,42**	0,37	0,27
-0,03 - 0,03	0,28	**0,35**	0,28	0,17
-0,02 - 0,02	0,24	**0,31**	0,22	0,12
-0,01 - 0,01	0,17	**0,26**	0,15	0,05
0 (kurtosis)	0,10	**0,21**	0,07	0,02

All years

Interval	Rosstat	**Self-reported**	Fed. ministries	FSO
-0,05 - 0,05	0,36	**0,40**	0,36	0,31
-0,03 - 0,03	0,27	**0,34**	0,27	0,22
-0,02 - 0,02	0,22	**0,31**	0,21	0,16
-0,01 - 0,01	0,16	**0,26**	0,14	0,11
0 (kurtosis)	0,09	**0,21**	0,07	0,05

Figure 34. Aggregate share of observations clustered around zero



Appendix 6. Indicators reported by regional governments (Decree 825)

Russian title	English title	Affiliation
1. Объем валового регионального продукта	Gross regional product	Rosstat (Fed.stat.bureau)
1.1. Объем прямых иностранных инвестиций в расчете на 1 жителя субъекта Российской Федерации	Amount of direct foreign investments per capita	Rosstat (Fed.stat.bureau)
1.2. Объем внешнеторгового оборота	Foreign trade turnover	Rosstat (Fed.stat.bureau)
1.3. Объем валового регионального продукта в расчете на 1 жителя субъекта Российской Федерации	Gross regional product per capita	Rosstat (Fed.stat.bureau)
10. Среднемесячная номинальная начисленная заработная плата работников, занятых в сфере экономики региона	Average monthly salary of workers in the region	Rosstat (Fed.stat.bureau)
101. Количество государственных (муниципальных) амбулаторно-поликлинических учреждений, финансирование которых осуществляется по результатам деятельности на основании подушевого норматива на прикрепленное население	Number of medical clinics financed through per capita standards	Regional administration
102. Количество государственных (муниципальных) больничных учреждений, финансирование которых осуществляется по результатам деятельности по законченному случаю	Number of medical clinics financed on the basis of cases closed	Regional administration
103. Доля государственных (муниципальных) учреждений здравоохранения, применяющих стандарты оказания медицинской помощи, в общем количестве государственных (муниципальных) учреждений здравоохранения, за исключением учреждений здравоохранения особого типа	The share of medical clinics using standards of service provision in the total number of medical clinics	Federal ministry
105. Численность лиц, систематически занимающихся физической культурой и спортом	The number of people systematically doing sports	Federal ministry
107. Обеспеченность спортивными сооружениями в субъекте Российской Федерации	Availability of sport facilities in the region	Federal ministry
108. Обеспеченность спортивными залами	Availability of sport halls	Federal ministry
109. Обеспеченность плоскостными спортивными сооружениями	Availability of stadiums	Federal ministry

Russian title	English title	Affiliation
11. Среднемесячная номинальная начисленная заработная плата работников, занятых в сфере сельского хозяйства региона	Average monthly salary of workers in agriculture	Rosstat (Fed.stat.bureau)
110. Обеспеченность плавательными бассейнами	Availability of swimming pools	Federal ministry
111. Удовлетворенность населения условиями для занятия физической культурой и спортом	Citizen satisfaction of doing sports	Federal security service
112. Удельный вес населения, систематически занимающегося физической культурой и спортом	The share of population doing sports systematically	Federal ministry
113. Доля учащихся (общеобразовательных учреждений, образовательных учреждений начального профессионального образования, образовательных учреждений среднего профессионального образования), занимающихся физической культурой и спортом, в общей численности учащихся соответствующих учреждений	The share of pupils doing sports systematically	Federal ministry
113.1. Численность спортсменов субъекта Российской Федерации, включенных в составы спортивных сборных команд Российской Федерации	The number of sportsmen enlisted in Russian national teams	Federal ministry
113.2. Доля спортсменов, зачисленных в составы спортивных сборных команд Российской Федерации, в общем количестве спортсменов, занимающихся на этапе совершенствования спортивного мастерства и этапе высшего спортивного мастерства	The share of sportsmen enlisted in Russian national teams in the total number of sportsmen doing sports at master's level	Regional administration
114. Расходы консолидированного бюджета субъекта Российской Федерации на физическую культуру и спорт	Budget expenditure on sports	Federal ministry
116. Удельный вес населения, участвующего в платных культурно-досуговых мероприятиях, проводимых государственными (муниципальными) учреждениями культуры	The shar of population taking part in organised cultural events	Regional administration
117. Количество экземпляров новых поступлений в библиотечные фонды общедоступных библиотек на 1 тыс. человек населения	The number of new books received by libraries per capita	Federal ministry
118.1. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, не получивших аттестат о среднем (полном) образовании	The number of high school graduates who did not receive the education certificate	Federal ministry

Russian title	English title	Affiliation
119. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, участвовавших в едином государственном экзамене по русскому языку	The number of high school graduates taking part in the national exam in Russian language	Federal ministry
12. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) учреждений здравоохранения	Average monthly salary of medical personnel	Rosstat (Fed.stat.bureau)
120. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, сдавших единый государственный экзамен по русскому языку	The number of high school graduates that passed the national exam in Russian language	Federal ministry
121. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, участвовавших в едином государственном экзамене по математике	The number of high school graduates taking part in the national exam in Mathematics	Federal ministry
122. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, сдавших единый государственный экзамен по математике	The number of high school graduates that passed the national exam in Mathematics	Federal ministry
123. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, участвовавших в едином государственном экзамене в городской местности	The number of high school graduates in urban areas	Federal ministry
123.1. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, участвовавших в едином государственном экзамене	The number of high school graduates that took part in the national exam	Federal ministry
124. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, участвовавших в едином государственном экзамене в сельской местности	The number of high school graduates in rural areas	Federal ministry
126. Количество преступлений, совершенных несовершеннолетними или при их соучастии	The number of crimes committed by underaged offenders	Federal ministry
127. Удовлетворенность населения качеством общего образования	Citizen satisfaction in the quality high school education	Federal security service
128. Количество государственных (муниципальных) общеобразовательных учреждений, расположенных в городской местности	The number of high schools in urban areas	Federal ministry

Russian title	English title	Affiliation
128.1. Количество государственных (муниципальных) общеобразовательных учреждений	The number of high schools	Federal ministry
129. Количество государственных (муниципальных) общеобразовательных учреждений, расположенных в сельской местности	The number of high schools in rural areas	Federal ministry
13. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) учреждений образования	Average nominal salary of high school staff	Rosstat (Fed.stat.bureau)
130. Количество государственных (муниципальных) общеобразовательных учреждений, здания которых находятся в аварийном состоянии или требуют капитального ремонта	The number of high schools requireing renovation	Federal ministry
131.1. Доля государственных (муниципальных) образовательных учреждений с постоянным пребыванием детей, здания которых находятся в аварийном состоянии или требуют капитального ремонта, в общем количестве государственных (муниципальных) образовательных учреждений с постоянным пребыванием детей	The share of boarding schools requireing capital renovation in the total number of boarding schools	Federal ministry
132. Численность лиц (среднегодовая), обучающихся в государственных (муниципальных) общеобразовательных учреждениях, расположенных в сельской местности	The number of pupils in high schools in rural areas	Federal ministry
133. Численность лиц (среднегодовая), обучающихся в государственных (муниципальных) общеобразовательных учреждениях, расположенных в городской местности	The number of pupils in high schools in urban areas	Federal ministry
134. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, расположенных в городской местности	The number of high school graduates in urban areas	Federal ministry
134.1. Численность выпускников государственных (муниципальных) общеобразовательных учреждений	The number of high school graduates	Federal ministry
135. Численность выпускников государственных (муниципальных) общеобразовательных учреждений, расположенных в сельской местности	The number of high school graduates in rural areas	Federal ministry
136. Удельный вес детей первой и второй групп здоровья в общей численности обучающихся в государственных (муниципальных)	The share of kids qualified in 1 and 2 health categories in the total number of pupils	Federal ministry

Russian title	English title	Affiliation
общеобразовательных учреждениях		
137. Численность работников (среднегодовая) государственных (муниципальных) общеобразовательных учреждений, расположенных в городской местности	The number of staff of high schools in ubran areas	Federal ministry
138. Численность работников (среднегодовая) государственных (муниципальных) общеобразовательных учреждений, расположенных в сельской местности	The number of staff of high schools in rural areas	Federal ministry
139. Численность учителей (среднегодовая) государственных (муниципальных) общеобразовательных учреждений, расположенных в городской местности	The number of teachers high schools in ubran areas	Federal ministry
14. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) учреждений социальной защиты населения	Average nominal salary of social services staff	Rosstat (Fed.stat.bureau)
140. Численность учителей (среднегодовая) государственных (муниципальных) общеобразовательных учреждений, расположенных в сельской местности	The number of teachers in high schools in rural areas	Federal ministry
141. Доля учителей государственных (муниципальных) общеобразовательных учреждений, имеющих стаж педагогической работы до 5 лет, в общей численности учителей государственных (муниципальных) общеобразовательных учреждений	The share of teachers with over 5 years of experience in the total number of teachers in high schools	Federal ministry
142. Численность прочего персонала (среднегодовая) (административно-управленческого, учебно-вспомогательного, младшего обслуживающего персонала, а также педагогических работников, не осуществляющих учебный процесс) государственных (муниципальных) общеобразовательных учреждений, расположенных в городской местности	The number of non-teaching staff in high schools in urban areas	Federal ministry
143. Численность прочего персонала (среднегодовая) (административно-управленческого, учебно-вспомогательного, младшего обслуживающего персонала, а также педагогических работников, не осуществляющих учебный процесс) государственных (муниципальных)	The number of non-teaching staff in high schools in rural areas	Federal ministry

Russian title	English title	Affiliation
общеобразовательных учреждений, расположенных в сельской местности		
144. Количество классов (среднегодовое) в государственных (муниципальных) общеобразовательных учреждениях, расположенных в городской местности	The number of classes in high shools in urban areas	Federal ministry
145. Количество классов (среднегодовое) в государственных (муниципальных) общеобразовательных учреждениях, расположенных в сельской местности	The number of classes in high schools in rural areas	Federal ministry
146. Средняя стоимость содержания одного класса в государственных (муниципальных) общеобразовательных учреждениях в субъекте Российской Федерации	Average cost of one class in high schools in the region	Regional administration
147. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) общеобразовательных учреждений	Average nominal monthly salary of staff of state owned high schools	Rosstat (Fed.stat.bureau)
148. Среднемесячная номинальная начисленная заработная плата учителей государственных (муниципальных) общеобразовательных учреждений	Average nominal salary of teachers in state owned high schools	Federal ministry
149. Среднемесячная номинальная начисленная заработная плата прочего персонала (административно-управленческого, учебно-вспомогательного, младшего обслуживающего персонала, а также педагогических работников, не осуществляющих учебный процесс) государственных (муниципальных) общеобразовательных учреждений	Average nominal salary of non-teaching staff in state owned high schools	Federal ministry
15. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) учреждений физической культуры и спорта	Average nominal salary of staff of sport establishments	Rosstat (Fed.stat.bureau)
150. Расходы консолидированного бюджета субъекта Российской Федерации на общее образование	Budget expenditure on education (high school level)	Federal ministry

Russian title	English title	Affiliation
151. Расходы консолидированного бюджета субъекта Российской Федерации на общее образование в части увеличения стоимости основных средств	Budget expenditure on education (high school level) in capital expenditure	Federal ministry
152. Расходы консолидированного бюджета субъекта Российской Федерации на общее образование в части текущих расходов	Budget expenditure on education (high school level) in current expenditure	Federal ministry
153. Расходы консолидированного бюджета субъекта Российской Федерации на общее образование в части текущих расходов на оплату труда и начислений на оплату труда	Budget expenditure on education (high school level) (salaries)	Federal ministry
154. Количество государственных (муниципальных) общеобразовательных учреждений, переведенных на нормативное подушевое финансирование	The number of high schools on per head financing	Federal ministry
155. Количество государственных (муниципальных) общеобразовательных учреждений, переведенных на новую (отраслевую) систему оплаты труда, ориентированную на результат	The number of high schools on new payment systems (performance pay)	Federal ministry
156. Доля детей, оставшихся без попечения родителей	The number of orphans	Federal ministry
157. Доля детей, оставшихся без попечения родителей, переданных на воспитание в семьи граждан Российской Федерации, постоянно проживающих на территории Российской Федерации (на усыновление (удочерение) и под опеку (попечительство)), в том числе по договору о приемной семье либо в случаях, предусмотренных законами субъектов Российской Федерации, по договору о патронатной семье (патронате, патронатном воспитании)	The number of fostered orphans	Federal ministry
158. Численность выпускников, завершивших обучение в отчетном году по образовательным программам начального профессионального образования за счет средств субъекта Российской Федерации	The number of state funded vocational colleges graduates	Federal ministry
159. Численность выпускников, завершивших обучение в отчетном году по образовательным программам начального профессионального образования за счет средств субъекта Российской Федерации, состоящих на регистрационном учете в качестве безработных	The number of state funded vocational colleges graduates registered as unemployed	#Н/Д

Russian title	English title	Affiliation
16. Среднемесячная номинальная начисленная заработная плата работников государственных (муниципальных) учреждений культуры и искусства	Average nominal monthly salary of staff of cultural establishments	Rosstat (Fed.stat.bureau)
160. Численность выпускников, завершивших обучение в отчетном году по образовательным программам среднего профессионального образования за счет средств субъекта Российской Федерации	The number of state funded graduates of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
161. Численность выпускников, завершивших обучение в отчетном году по образовательным программам среднего профессионального образования за счет средств субъекта Российской Федерации, состоящих на регистрационном учете в качестве безработных	The number of state funded graduates of middle-tier vocational colleges registered as unemployed	#Н/Д
166. Количество преступлений, совершенных несовершеннолетними, обучающимися в государственных образовательных учреждениях начального профессионального образования, или при их соучастии	The number of crimes committed by students of entry-level vocational colleges	Federal ministry
167. Количество преступлений, совершенных несовершеннолетними, обучающимися в государственных образовательных учреждениях среднего профессионального образования, или при их соучастии	The number of crimes committed by students of middle-level vocational colleges	Federal ministry
168. Количество государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации	The number of entry-level vocational colleges	Federal ministry
169. Количество государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации, здания которых находятся в аварийном состоянии или требуют капитального ремонта	The number of entry-level vocational colleges with buildings requireing capital renovation	#Н/Д
17. Доля среднесписочной численности работников (без внешних совместителей) малых и средних предприятий в среднесписочной численности работников (без внешних совместителей) всех предприятий и организаций	The share employees of small and medium sided businesess in the total number of emplyees	Rosstat (Fed.stat.bureau)
170. Численность лиц, обучающихся в государственных образовательных учреждениях начального профессионального образования субъекта Российской Федерации	The number of students of state-run entrly-level vocational colleges	Federal ministry

Russian title	English title	Affiliation
170.1. Численность лиц, обучающихся по образовательным программам начального профессионального образования за счет средств субъекта Российской Федерации	The number of state funded students of entry-level vocational colleges	Federal ministry
170.2. Доля лиц, обучающихся по образовательным программам начального профессионального образования за счет средств субъекта Российской Федерации в учреждениях среднего и высшего профессионального образования, в общей численности лиц, обучающихся по образовательным программам начального профессионального образования за счет средств субъекта Российской Федерации	The share of state funded students in the total number of students at middle-tier vocational colleges	Federal ministry
172. Численность лиц, обучающихся в государственных образовательных учреждениях начального профессионального образования субъекта Российской Федерации за счет внебюджетных средств	The number of students of state-run entrly-level vocational colleges funded from off-budget sources	#Н/Д
173. Численность выпускников, завершивших обучение по образовательным программам начального профессионального образования в отчетном году за счет средств субъекта Российской Федерации, получивших направление на работу в организации	The number of graduates of entry-level vocational colleges who received a work placement	#Н/Д
174. Численность работников государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации (физические лица)	The number of staff of entry-level vocational colleges	Federal ministry
175. Численность преподавателей государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации (физические лица)	The number of teachers at entrly-level vocational colleges	Federal ministry
176. Численность мастеров производственного обучения государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации (физические лица)	The number of instructors at entry-level vocational colleges	Federal ministry

Russian title	English title	Affiliation
178. Численность прочего персонала государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации (административно-управленческого, учебно-вспомогательного, младшего обслуживающего персонала, а также педагогических работников, не осуществляющих учебный процесс) (физические лица)	The number of other staff at entry-level vocational colleges	Federal ministry
179. Доля преподавателей государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации со стажем работы до 5 лет в общей численности преподавателей государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации	The share of teachers with over 5 years of experience in the total number of teachers at entry-level vocational colleges	Federal ministry
18. Количество малых и средних предприятий в расчете на 1 тыс. человек населения субъекта Российской Федерации	The number of small and medium sized business per 1000 people	Rosstat (Fed.stat.bureau)
180. Среднемесячная номинальная начисленная заработная плата преподавателей государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации	Average nominal monthly salary of teachers of entry-level vocational colleges	#Н/Д
181. Среднемесячная номинальная начисленная заработная плата мастеров производственного обучения государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации	Average nominal montly income of instructors of entry-level vocational colleges	Federal ministry
182. Среднемесячная номинальная начисленная заработная плата работников государственных образовательных учреждений начального профессионального образования	Average nominal monthly income of staff of entry-level vocational colleges	#Н/Д
183. Расходы консолидированного бюджета субъекта Российской Федерации на начальное профессиональное образование	Budget expediture on entry-level vocational education	Federal ministry
184. Расходы консолидированного бюджета субъекта Российской Федерации на начальное профессиональное образование в части увеличения стоимости основных средств	Budget expediture on entry-level vocational education - capital expenditure	Federal ministry

Russian title	English title	Affiliation
185. Расходы консолидированного бюджета субъекта Российской Федерации на начальное профессиональное образование в части текущих расходов	Budget expediture on entry-level vocational education - current expenditure	Federal ministry
186. Расходы консолидированного бюджета субъекта Российской Федерации на начальное профессиональное образование в части текущих расходов на оплату труда и начислений на оплату труда	Budget expediture on entry-level vocational education - salaries	Federal ministry
187. Количество государственных образовательных учреждений начального профессионального образования субъекта Российской Федерации, переведенных на новую (отраслевую) систему оплаты труда, ориентированную на результат	The number of entry-level vocational colleges where performance pay has been implemented	Regional administration
188. Количество государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации	The number of middle-level vocational colleges	Rosstat (Fed.stat.bureau)
189. Количество государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации, здания которых находятся в аварийном состоянии или требуют капитального ремонта	The number of middle-level vocational colleges with buildings requiring capital renovation	Rosstat (Fed.stat.bureau)
19. Доля продукции, произведенной малыми предприятиями, в общем объеме валового регионального продукта	The share of small businesses in gross regional product	Rosstat (Fed.stat.bureau)
190. Численность лиц, обучающихся в государственных образовательных учреждениях среднего профессионального образования субъекта Российской Федерации	The number of students at middle-level vocational colleges	Rosstat (Fed.stat.bureau)
190.1. Численность лиц, обучающихся по образовательным программам среднего профессионального образования за счет средств субъекта Российской Федерации	The number of students at middle-level vocational colleges (state funded)	Rosstat (Fed.stat.bureau)
192. Численность лиц, обучающихся в государственных образовательных учреждениях среднего профессионального образования субъекта Российской Федерации за счет внебюджетных средств	The number of students at middle-level vocational colleges (funded from off-budget sources)	Regional administration
193. Численность выпускников, завершивших обучение в отчетном году по образовательным программам среднего профессионального образования за счет средств субъекта Российской Федерации,	The number of state funded graduates of middle-tier vocational colleges	#Н/Д

Russian title	English title	Affiliation
получивших направление на работу в организации		
194. Численность работников государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации (физические лица)	The number of staff of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
195. Численность преподавателей государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации (физические лица)	The number of instructors of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
198. Численность прочего персонала государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации (административно-управленческого, учебно-вспомогательного, младшего обслуживающего персонала, а также педагогических работников, не осуществляющих учебный процесс) (физические лица)	The number of other staff at middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
199. Доля преподавателей государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации со стажем работы до 5 лет в общей численности преподавателей государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации	The share of teachers with more than 5 years of experience in the overall number of teachers of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
2.1. Индекс промышленного производства	Industrial production index	Rosstat (Fed.stat.bureau)
20. Удельный вес прибыльных крупных и средних сельскохозяйственных организаций в их общем числе	The share of profitable large and medium sized businesses	Rosstat (Fed.stat.bureau)
200. Среднемесячная номинальная начисленная заработная плата преподавателей государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации	Average nominal monthly salary of teachers of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)
202. Среднемесячная номинальная начисленная заработная плата работников государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации	Average nominal monthly salary of staff of middle-tier vocational colleges	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
203. Расходы консолидированного бюджета субъекта Российской Федерации на среднее профессиональное образование	Budget expenditure on education (middle tier vocational colleges)	Federal ministry
204. Расходы консолидированного бюджета субъекта Российской Федерации на среднее профессиональное образование в части увеличения стоимости основных средств	Budget expenditure on education (middle tier vocational colleges) - capital expenditure	Federal ministry
205. Расходы консолидированного бюджета субъекта Российской Федерации на среднее профессиональное образование в части текущих расходов	Budget expenditure on education (middle tier vocational colleges) - current expenditure	Federal ministry
206. Расходы консолидированного бюджета субъекта Российской Федерации на среднее профессиональное образование в части текущих расходов на оплату труда и начислений на оплату труда	Budget expenditure on education (middle tier vocational colleges) - salaries	Federal ministry
207. Количество государственных образовательных учреждений среднего профессионального образования субъекта Российской Федерации, переведенных на новую (отраслевую) систему оплаты труда, ориентированную на результат	The number of middle-tier vocational colleges where performance pay has been implemented	Regional administration
208. Общая площадь жилых помещений, приходящаяся в среднем на 1 жителя субъекта Российской Федерации	Housing area per capita	Rosstat (Fed.stat.bureau)
209. Количество жилых помещений (квартир) в расчете на 1 тыс. человек населения	Number of flats per 1000 people	Rosstat (Fed.stat.bureau)
21. Расходы консолидированного бюджета субъекта Российской Федерации на сельское хозяйство в расчете на 1 рубль произведенной сельскохозяйственной продукции	Budget expenditure on agriculture per 1 rouble of agricultural production	Federal ministry
210. Соотношение средней рыночной стоимости стандартной квартиры общей площадью 54 кв.м. и среднего годового совокупного денежного дохода семьи, состоящей из 3 человек	The ratio between the market value of a 54 sq.m flat (primary market) and gross annual income of a family of 3 members	Federal ministry
211. Отношение средней цены 1 кв. метра общей площади на первичном рынке жилья к среднедушевым доходам населения в субъекте Российской Федерации	The ratio between the price of 1 sq.m of housing (primary market) and average income of workers in the region	Rosstat (Fed.stat.bureau)
211.1. Отношение средней цены 1 кв. метра общей площади на вторичном рынке жилья к среднедушевым доходам населения в	The ratio between the price of 1 sq.m of housing (secondary market) and average income of workers	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
субъекте Российской Федерации	in the region	
213. Общая площадь жилых помещений, приходящаяся в среднем на 1 жителя субъекта Российской Федерации, введенная в действие за год	Total area of housing built annually per capita	Rosstat (Fed.stat.bureau)
214. Количество жилых помещений (квартир) в расчете на 1 тыс. человек населения, введенных в действие за год	The number of flats built annually per 1000 people	Rosstat (Fed.stat.bureau)
215. Общая площадь жилых помещений, строительство которых предусмотрено в соответствии с выданными разрешениями на строительство жилых зданий, в среднем на 1 жителя субъекта Российской Федерации	The overall area of housing licenced to be built per capita	Regional administration
217. Площадь земельных участков, предоставленных для жилищного строительства и комплексного освоения в целях жилищного строительства, в расчете на душу населения субъекта Российской Федерации	The area of land licenced for housing development per capita	Regional administration
218. Средняя продолжительность периода с даты подписания протокола о результатах аукционов по предоставлению земельных участков для жилищного строительства до получения разрешения на строительство	The average time between the tender for housing development and issue of construction lisence	Regional administration
219. Средняя продолжительность периода с даты выдачи разрешения на строительство жилого здания до даты получения разрешения на ввод жилого здания в эксплуатацию	The average time between issuing of construction permit and completion of development	Regional administration
22. Уровень занятости сельского населения трудоспособного возраста	Employment level in rural areas	Rosstat (Fed.stat.bureau)
220. Средняя продолжительность периода с даты подачи заявки на предоставление земельного участка для строительства до даты получения разрешения на строительство	The average time between application for housing development permit and issue of the permit	Regional administration
221. Площадь земельных участков, предоставленных для жилищного строительства, в отношении которых с даты принятия решения о предоставлении земельного участка или подписания протокола о результатах торгов (конкурсов, аукционов) не было получено разрешение на ввод в эксплуатацию в течение 3 лет	The area of land sites for which the time between issuing the development license and completion of construction exeeded 3 years	Regional administration
221.1. Объем приостановленного жилищного строительства	The amount of frozen housing development	Federal ministry

Russian title	English title	Affiliation
222. Потери организаций коммунального комплекса вследствие перерасчета платежей потребителей из-за предоставления коммунальных ресурсов и услуг ненадлежащего качества и (или) с перерывами, превышающими установленную продолжительность, в расчете на единицу доходов от реализации услуг по основному виду деятельности	Communication companies' losses in subsidies of tariffs	Regional administration
224. Удовлетворенность населения жилищно-коммунальными услугами	Citizen satisfaction of housing communication services	Federal security service
225. Доля населения, обеспеченного питьевой водой, отвечающей требованиям безопасности, в общей численности населения субъекта Российской Федерации	The share of population supplied with potable water	Federal ministry
226. Доля утечек и неучтенного расхода воды в суммарном объеме воды, поданной в сеть	The share of leakage in overall water supply	Rosstat (Fed.stat.bureau)
227. Доля потерь тепловой энергии в суммарном объеме отпуска тепловой энергии	The share of lost thermal enere in overall supply	Rosstat (Fed.stat.bureau)
233. Уровень износа коммунальной инфраструктуры	The level of tear and wear of communication infrastructure	Regional administration
234. Доля многоквартирных жилых домов с износом более 31 процента, в которых проведен капитальный ремонт, в общем количестве многоквартирных жилых домов, требующих капитального ремонта	The share of housing with over 31% wear and tear	Regional administration
235. Доля населения, проживающего в многоквартирных домах, признанных в установленном порядке аварийными	The share of population living in houses requiring capital renovation	Rosstat (Fed.stat.bureau)
236. Расходы консолидированного бюджета субъекта Российской Федерации на финансирование жилищно-коммунального хозяйства	Budget expenditure on communal services	Federal ministry
237. Расходы консолидированного бюджета субъекта Российской Федерации на финансирование жилищно-коммунального хозяйства в части увеличения стоимости основных средств	Budget expenditure on communal services - capital expenditure	Federal ministry

Russian title	English title	Affiliation
238. Расходы консолидированного бюджета субъекта Российской Федерации на финансирование жилищно-коммунального хозяйства в части компенсации разницы между экономически обоснованными тарифами и тарифами, установленными для населения, и покрытия убытков, возникших в связи с применением регулируемых цен на жилищно-коммунальные услуги	Budget expenditure on communal services - subsidies of tariffs	Federal ministry
239. Уровень возмещения населением затрат за предоставление жилищно-коммунальных услуг по установленным для населения тарифам	The share of costs of communal services covered by user charges	Rosstat (Fed.stat.bureau)
24. Доля обрабатываемой пашни в общей площади пашни	The share of tilled ploughed field	Federal ministry
240. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют непосредственное управление собственниками помещений в многоквартирном доме	The share of appartment blocks in which owners chose and operate direct owner's control scheme over communal areas	Federal ministry
241. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют управление многоквартирными домами посредством товариществ собственников жилья либо жилищных кооперативов или иного специализированного потребительского кооператива	The share of appartment blocks in which owners chose and operate one of available schemes of control over communal areas	Federal ministry
242. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации	The share of appartment blocks in which owners chose an operator over communal areas	Federal ministry
243. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации муниципальной формы собственности	The share of appartment blocks in which owners chose a municipal operator over communal areas	Federal ministry
244. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации муниципальной формы собственности в форме муниципальных учреждений	The share of appartment blocks in which owners chose a municipal operator of municipal organisation	Federal ministry

Russian title	English title	Affiliation
245. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации государственной формы собственности	The share of appartment blocks in which owners chose a state owned company as an operator of communal areas and infrastructure	Federal ministry
246. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации государственной формы собственности в форме государственных учреждений	The share of appartment blocks in which owners chose a state owned company as an operator of communal areas and infrastructure (in the form of state owned company)	Federal ministry
247. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством управляющей организации частной формы собственности	The share of appartment blocks in which owners chose an operator of communal areas (in the form of privately owned company)	Federal ministry
248. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством хозяйственных обществ со 100-процентной долей, находящейся в муниципальной или государственной собственности	The share of appartment blocks in which owners chose an operator of communal areas (in the form of a joint stock company with 100% municipal and state ownership)	Federal ministry
249. Доля многоквартирных домов в целом по субъекту Российской Федерации, в которых собственники помещений выбрали и реализуют способ управления многоквартирными домами посредством хозяйственных обществ с долей участия, не превышающей 25%, находящейся в государственной (муниципальной) собственности	The share of appartment blocks in which owners chose an operator of communal areas (in the form of a joint stock company with no more than 25% municipal and state ownership)	Federal ministry
25. Общий объем мощности, заявленной для технологического присоединения к объектам электросетевого хозяйства в отчетном году	Overall volume of electrical capacity in objects connected to the grid in the current year	Regional administration

Russian title	English title	Affiliation
250. Доля организаций, осуществляющих управление многоквартирными домами и оказание услуг по содержанию и ремонту общего имущества в многоквартирных домах, с долей участия в уставном капитале субъектов Российской Федерации и муниципальных образований не более чем 25 процентов в общем количестве организаций, осуществляющих свою деятельность на территории муниципального образования (территориях субъектов Российской Федерации - городов федерального значения Москвы и Санкт-Петербурга) и управление многоквартирными домами (кроме товариществ собственников жилья, жилищных, жилищно-строительных кооперативов или иных специализированных потребительских кооперативов)	The share of appartment blocks in which owners chose an operator of communal areas (in the form of a joint stock company with no more than 25% municipal and state ownership) in the total number of appartment blocks that have chosen an operator	Federal ministry
251. Доля организаций коммунального комплекса с долей участия в уставном капитале субъектов Российской Федерации и (или) муниципальных образований не более чем 25 процентов, осуществляющих производство товаров, оказание услуг по электро-, газо-, тепло- и водоснабжению, водоотведению, очистке сточных вод, а также эксплуатацию объектов для утилизации (захоронения) твердых бытовых отходов, использующих объекты коммунальной инфраструктуры на праве частной собственности, по договору аренды или концессионному соглашению	The share of organisations with less than 25% municipal and state stock working as concessions or public private partnerships	Federal ministry
252. Доля доходов от реализации услуг по основному виду деятельности убыточных организаций коммунального комплекса в общем объеме доходов от реализации услуг по основному виду деятельности организаций коммунального комплекса в субъекте Российской Федерации	The share of income generated by non-profitable organisations of communal services in the overall income of organisations of communal services	Rosstat (Fed.stat.bureau)
253. Отношение финансового результата от реализации услуг по основному виду деятельности убыточных организаций коммунального комплекса к доходам от реализации услуг по основному виду деятельности убыточных организаций коммунального комплекса в субъекте Российской Федерации	The ration between losses of non-profitable communal service providers and overall income of communal service providers	Rosstat (Fed.stat.bureau)
254. Доля убыточных организаций жилищно-коммунального хозяйства	The share of non-profitable communal service	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
	providers	
255. Доля муниципальных образований, в которых тарифы на холодное водоснабжение для различных групп потребителей коммунальных услуг установлены без учета необходимости покрытия затрат на предоставление соответствующего вида коммунальных услуг одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве затрат на предоставление холодного водоснабжения одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве муниципальных образований субъекта Российской Федерации	The share of municipalities where prices of cold water do not take into accout cross-subsidies for different types of users	Regional administration
256. Доля муниципальных образований, в которых тарифы на горячее водоснабжение для различных групп потребителей коммунальных услуг установлены без учета необходимости покрытия затрат на предоставление соответствующего вида коммунальных услуг одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве затрат на предоставление горячего водоснабжения одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве муниципальных образований субъекта Российской Федерации	The share of municipalities where prices of hot water do not take into accout cross-subsidies for different types of users	Regional administration
257. Доля муниципальных образований, в которых тарифы на водоотведение и очистку сточных вод для различных групп потребителей коммунальных услуг установлены без учета необходимости покрытия затрат на предоставление соответствующего вида коммунальных услуг одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве затрат на предоставление водоотведения и очистки сточных вод одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве муниципальных образований субъекта Российской Федерации	The share of municipalities where prices of sewage do not take into accout cross-subsidies for different types of users	Regional administration

Russian title	English title	Affiliation
258. Доля муниципальных образований, в которых тарифы на тепловую энергию для различных групп потребителей коммунальных услуг установлены без учета необходимости покрытия затрат на предоставление соответствующего вида коммунальных услуг одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве затрат на предоставление тепловой энергии одной группе потребителей за счет тарифов, установленных для другой группы потребителей, в общем количестве муниципальных образований субъекта Российской Федерации	The share of municipalities where prices of heating do not take into accout cross-subsidies for different types of users	Regional administration
26. Заявленная мощность, которая не была удовлетворена в связи с отсутствием технической возможности технологического присоединения к объектам электросетевого хозяйства	Declared electrical capacity that was not met due to technical difficulties with connecting objects to the grid	Regional administration
260. Доля муниципальных образований, в которых предоставление установленных федеральными законами и законами субъектов Российской Федерации мер социальной поддержки гражданам на оплату жилого помещения и коммунальных услуг (за исключением субсидий гражданам на оплату жилого помещения и коммунальных услуг) осуществляется в денежной форме (в том числе путем перечисления средств на предоставление таких мер через банковские счета в банках, организации связи или иным способом)	The share of municipalities where subsidies to low-income groups on housing are given out in monetary form	Rosstat (Fed.stat.bureau)
261. Доля семей, получающих жилищные субсидии на оплату жилого помещения и коммунальных услуг, в общем количестве семей в субъекте Российской Федерации	The share of families receiving housing subsidies in the total number of families in the region	Rosstat (Fed.stat.bureau)
262. Доля многоквартирных домов, расположенных на земельных участках, в отношении которых осуществлен государственный кадастровый учет	The share of appartment blocks located on land sites that have been included in the cadastre	Regional administration
263. Доля протяженности автомобильных дорог общего пользования регионального значения, не отвечающих нормативным требованиям, в общей протяженности автомобильных дорог общего пользования регионального значения	The share of low quality regional roads in the total milege of regional roads	Regional administration

Russian title	English title	Affiliation
264. Доля протяженности автомобильных дорог общего пользования местного значения, не отвечающих нормативным требованиям, в общей протяженности автомобильных дорог общего пользования местного значения	The share of local low quality roads in the total milege of local roads	Rosstat (Fed.stat.bureau)
265. Доля протяженности автомобильных дорог общего пользования регионального или межмуниципального значения, работающих в режиме перегрузки, в общей протяженности автомобильных дорог общего пользования регионального или межмуниципального значения	The share of overcrowded regional and inter-municipal roads in the total milege of regional and inter-municipal roads	Regional administration
266. Доля дорожно-транспортных происшествий, совершению которых сопутствовало наличие неудовлетворительных дорожных условий, в общем количестве дорожно-транспортных происшествий	The share of road accidents caused by (or accompanied with) low quality of road paving	Federal ministry
267. Доля протяженности автомобильных дорог общего пользования регионального значения, содержание которых в отчетном году осуществляется в соответствии с государственными долгосрочными контрактами, заключенными с организациями негосударственной и немуниципальной форм собственности, в общей протяженности автомобильных дорог общего пользования регионального значения	The share of regional roads maintained by non-municipal and non-public companies in accordance with long-term contracts in the total milege of regional roads	Regional administration
268. Доля протяженности автомобильных дорог общего пользования местного значения, содержание которых в отчетном году осуществляется в соответствии с муниципальными долгосрочными контрактами, заключенными с организациями негосударственной и немуниципальной форм собственности, в общей протяженности автомобильных дорог общего пользования местного значения	The share of municipal roads maintained by non-municipal and non-public companies in accordance with long-term contracts in the total milege of municipal roads	Regional administration
269. Доля лиц, ранее осуждавшихся за совершение преступлений, в общем количестве лиц, осужденных на основании обвинительных приговоров, вступивших в законную силу	The share of repeating offenders in the total number of criminal offenders	Regional administration
27. Ставка платы за технологическое присоединение к электрическим сетям на уровне напряжения ниже 35 кВ и мощности менее 10 000 кВА	Cost of connecting to the electrical grid (below 35 kwatts and 10 000 kva)	Regional administration
270. Количество зарегистрированных преступлений	The number of registered crimes	Federal ministry
271. Оценка населением уровня криминогенности в субъекте Российской Федерации	Citizens' assessment of the criminal situation in the region	Federal security service

Russian title	English title	Affiliation
272. Удовлетворенность населения деятельностью органов исполнительной власти субъекта Российской Федерации по обеспечению безопасности граждан	Citizen satisfaction of the work of regional government	Federal security service
273. Расходы консолидированного бюджета субъекта Российской Федерации на правоохранительную деятельность	Budgetary expenditure on policing	Federal ministry
274. Объем просроченной кредиторской задолженности государственных (муниципальных) учреждений	Total overdue debts of public sector organisations	Federal ministry
275. Задолженность бюджета субъекта Российской Федерации и бюджетов муниципальных образований по исполнению обязательств перед гражданами	The amount of overdue payments to citizens	Federal ministry
276. Объем незавершенного в установленные сроки строительства, осуществляемого за счет средств консолидированного бюджета субъекта Российской Федерации	The share of unfinished construction financed by the regional budget	Regional administration
277. Доля налоговых и неналоговых доходов консолидированного бюджета субъекта Российской Федерации в общем объеме доходов консолидированного бюджета субъекта Российской Федерации (без учета субвенций)	The share of tax and non-tax incomes in the total amount of income of regional budget (excluding subventions)	Federal ministry
278. Доля налоговых и неналоговых доходов бюджетов муниципальных районов в общем объеме доходов бюджетов муниципальных районов (без учета субвенций)	The share of tax and non-tax incomes in the total amount of income of municipal budgets (excluding subventions)	Federal ministry
279. Доля налоговых и неналоговых доходов бюджетов городских округов в общем объеме доходов бюджетов городских округов (без учета субвенций)	The share of tax and non-tax incomes in the total amount of income of urban districts (excluding subventions)	Federal ministry
28. Количество планируемых к вводу в эксплуатацию в соответствии с утвержденными инвестиционными программами объектов электросетевого хозяйства	The number of power generting facilities planned to be finished in the current year	Regional administration
280. Доля доходов бюджетов муниципальных образований (без учета субвенций) в общем объеме доходов консолидированного бюджета субъекта Российской Федерации	The share of revenues of municipal budgets in the total amount of revenues of the regional budget	Federal ministry

Russian title	English title	Affiliation
280.1. Доля налоговых и неналоговых доходов бюджетов муниципальных образований в общем объеме налоговых и неналоговых доходов консолидированного бюджета субъекта Российской Федерации	The share of tax and non-tax revenues of municipal budgets in the total amount of tax and non-tax revenues of regional budget	Federal ministry
281. Доля расходов консолидированного бюджета субъекта Российской Федерации на увеличение стоимости основных средств в общем объеме расходов консолидированного бюджета субъекта Российской Федерации	The share of capital expenditure of regional budget	Federal ministry
282. Удовлетворенность населения деятельностью органов исполнительной власти субъекта Российской Федерации	Citizen satisfaction with the work of regional administration	Federal security service
282.1. в том числе их информационной открытостью	Citizen satisfaction with transparency of the work of regional administration	Federal security service
283. Количество унитарных и казенных предприятий в субъекте Российской Федерации	The number of unitary and treasury-owned enterprises in the region	Rosstat (Fed.stat.bureau)
284. Количество государственных унитарных и казенных предприятий в субъекте Российской Федерации	The number of state-owned unitary and treasury-owned enterprises in the region	Rosstat (Fed.stat.bureau)
285. Количество муниципальных унитарных и казенных предприятий в субъекте Российской Федерации	The number of municipal-owned unitary and treasury-owned enterprises in the region	Rosstat (Fed.stat.bureau)
285.1. Количество государственных (муниципальных) услуг, предоставляемых органами исполнительной власти субъекта Российской Федерации (органами местного самоуправления), учреждениями субъекта Российской Федерации (муниципальными учреждениями) в электронном виде	The number of public services provided online	Federal ministry
285.2. Количество государственных (муниципальных) услуг, предоставляемых органами исполнительной власти субъекта Российской Федерации (органами местного самоуправления), учреждениями субъекта Российской Федерации (муниципальными учреждениями)	The number of public services provided by regional and municipal organizations in the region	Federal ministry
285.3. Количество первоочередных государственных (муниципальных) услуг, предоставляемых органами исполнительной власти субъекта Российской Федерации (органами местного самоуправления) и учреждениями субъекта Российской Федерации (муниципальными	The number of first-priority public services provided online	Federal ministry

Russian title	English title	Affiliation
учреждениями) в электронном виде		
286. Среднемесячная начисленная заработная плата гражданских служащих органов исполнительной власти субъекта Российской Федерации	Average monthly salary of civil servants working at regional public sector organizations	Rosstat (Fed.stat.bureau)
287. Численность лиц, занятых в органах исполнительной власти субъекта Российской Федерации	The number of staff at regional public sector bureaus	Rosstat (Fed.stat.bureau)
288. Численность государственных служащих в органах исполнительной власти субъекта Российской Федерации	The number of civil servants at regional public sector bureaus	Rosstat (Fed.stat.bureau)
289. Численность муниципальных служащих в органах местного самоуправления субъекта Российской Федерации	The number of municipal civil servants at municipal agencies	Rosstat (Fed.stat.bureau)
29. Количество введенных в эксплуатацию в соответствии с утвержденными инвестиционными программами объектов электросетевого хозяйства	The number of launched electricity generating objects	Regional administration
290. Доля расходов консолидированного бюджета субъекта Российской Федерации, формируемых в рамках программ, в общем объеме расходов консолидированного бюджета субъекта Российской Федерации (без учета субвенций на исполнение делегируемых полномочий)	The share of expenditure done through public sector programmes in the total expenditure of the regional budget	Regional administration
291. Доля расходов консолидированного бюджета субъекта Российской Федерации на финансирование услуг социальной сферы, оказываемых автономными учреждениями и негосударственными (немуниципальными) организациями, в общем объеме расходов консолидированного бюджета субъекта Российской Федерации на финансирование отраслей социальной сферы	The share of expenditure on social security done through autonomous and non-for-profit organisations in the total expenditure on social security	Regional administration

Russian title	English title	Affiliation
292. Доля расходов консолидированного бюджета субъекта Российской Федерации на оказание бюджетных услуг, оказываемых негосударственными (немуниципальными) организациями, в общем объеме расходов консолидированного бюджета субъекта Российской Федерации на финансирование бюджетных услуг	The share of expenditure of the regional budget on public services done through non-public and non-municipal organisationas in the total expenditure on public services	Regional administration
293. Расходы консолидированного бюджета субъекта Российской Федерации	Expenditure of the consolidated regional budget	Federal ministry
294. Расходы консолидированного бюджета субъекта Российской Федерации в части увеличения стоимости основных средств	Expenditure of the consolidated regional budget (capital expenditure)	Federal ministry
295. Расходы консолидированного бюджета субъекта Российской Федерации в части расходов на содержание работников органов государственной власти и органов местного самоуправления	Expenditure of the consolidated regional budget (staff payroll)	Federal ministry
295.1. Доля стоимости государственных (муниципальных) контрактов, заключенных по результатам несостоявшихся торгов и запросов котировок у единственного поставщика (исполнителя, подрядчика), в общей стоимости заключенных государственных (муниципальных) контрактов	The share of public contracts procured via unsuccessful tenders or tenders with a single applicant	Rosstat (Fed.stat.bureau)
295.2. Доля государственных (муниципальных) контрактов, заключенных путем запроса котировок, либо по результатам несостоявшихся торгов и запросов котировок у единственного поставщика (исполнителя, подрядчика), либо закупок малого объема в общем количестве заключенных государственных (муниципальных) контрактов, за исключением государственных (муниципальных) контрактов, заключенных с единственным поставщиком (исполнителем, подрядчиком) без проведения торгов	The share of public and municipal contracts procured through quotations from a single supplier or through unsuccessful tenders or small scale contracts excluding contracts with a single supplier without tendering.	Rosstat (Fed.stat.bureau)
295.3. Среднее количество поставщиков, принявших участие в одном конкурсе, аукционе, закупке	The average number of suppliers in tenders	Regional administration
295.4. Доля стоимости государственных (муниципальных) контрактов, осуществленных посредством электронных аукционов, в общей стоимости государственных (муниципальных) контрактов	The share of public contracts procured through online tendering in the total value of public contracts	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
295.5. Отношение объема государственного долга субъекта Российской Федерации по состоянию на 1 января года, следующего за отчетным, к общему годовому объему доходов бюджета субъекта Российской Федерации в отчетном финансовом году (без учета объемов безвозмездных поступлений)	The ratio between regional debt and overall budget revenue	Federal ministry
296. Энергоемкость валового регионального продукта	Energy-output ration of regional gross regional product	Rosstat (Fed.stat.bureau)
297. Доля объема электрической энергии, расчеты за потребление которой осуществляются на основании показаний приборов учета, в общем объеме электрической энергии, потребляемой на территории субъекта Российской Федерации	The share of electrical power procured with meters in the total power supplied	Regional administration
298. Доля объема тепловой энергии, расчеты за потребление которой осуществляются на основании показаний приборов учета, в общем объеме тепловой энергии, потребляемой на территории субъекта Российской Федерации	The share of heat power procured with meters in the total heat power supplied	Regional administration
299. Доля объема горячей воды, расчеты за потребление которой осуществляются на основании показаний приборов учета, в общем объеме горячей воды, потребляемой на территории субъекта Российской Федерации	The share of hot water procured with meters in the total hot water procured	Regional administration
3. Объем инвестиций в основной капитал (за исключением бюджетных средств) в расчете на 1 человека	Capital investment per capita	Rosstat (Fed.stat.bureau)
30. Трансформаторная мощность, введенная в эксплуатацию в соответствии с утвержденными инвестиционными программами	Transformer power launched in accordance with investment programmes	Regional administration
300. Доля объема холодной воды, расчеты за потребление которой осуществляются на основании показаний приборов учета, в общем объеме холодной воды, потребляемой на территории субъекта Российской Федерации	The share of cold water procured with meters in the total cold water procured	Regional administration
301. Доля объема природного газа, расчеты за потребление которого осуществляются на основании показаний приборов учета, в общем объеме природного газа, потребляемого на территории субъекта Российской Федерации	The share of gas supplied with meters in the total gas supplied	Regional administration

Russian title	English title	Affiliation
302. Доля энергетических ресурсов, производимых с использованием возобновляемых источников энергии, в общем объеме энергетических ресурсов, производимых на территории субъекта Российской Федерации	The share of power generated from renewable sources in the total power generated	Rosstat (Fed.stat.bureau)
303. Удельная величина потребления электрической энергии в многоквартирных домах	Unit quantity of electrical power consumption in appartment blocks	Regional administration
304. Удельная величина потребления тепловой энергии в многоквартирных домах	Unit quantity of heat consumption in appartment blocks	Regional administration
305. Удельная величина потребления горячей воды в многоквартирных домах	Unit quantity of hot water consumption in appartment blocks	Regional administration
306. Удельная величина потребления холодной воды в многоквартирных домах	Unit quantity of cold water consumption in appartment blocks	Regional administration
307. Удельная величина потребления природного газа в многоквартирных домах	Unit quantity of gas consumption in appartment blocks	Regional administration
308. Расходы консолидированного бюджета субъекта Российской Федерации на реализацию региональной программы в области энергосбережения и повышения энергетической эффективности	Budget expenditure on programmes promoting energy efficiency	Federal ministry
309. Количество субъектов хозяйственной и иной деятельности с установленными нормативами предельно допустимых выбросов вредных (загрязняющих) веществ в атмосферный воздух, расположенных на территории субъекта Российской Федерации и подлежащих федеральному статистическому наблюдению по форме 2-ТП (воздух) «Сведения об охране атмосферного воздуха»	The number of facilities with air polution quotas monitored in accordance with statistical form 2-TP	Rosstat (Fed.stat.bureau)
31. Объем планируемой к вводу в эксплуатацию в соответствии с утвержденными инвестиционными программами трансформаторной мощности	Electrical transformer capacity planned to be launched this year in accordance with approved investment programmes	Regional administration
310. Общее количество субъектов хозяйственной и иной деятельности, расположенных на территории субъекта Российской Федерации и подлежащих федеральному статистическому наблюдению по форме 2-ТП (воздух) «Сведения об охране атмосферного воздуха»	The number of firms in the region monitored in accordance with statistical form 2-tp (Air polution)	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
311. Объем выбросов вредных (загрязняющих) веществ в атмосферный воздух от стационарных источников, расположенных на территории субъекта Российской Федерации	Amount of air polluting agents emitted by facilities located in the region	Rosstat (Fed.stat.bureau)
312. Объем выбросов вредных (загрязняющих) веществ в атмосферный воздух от автомобильного транспорта, зарегистрированного на территории субъекта Российской Федерации	The amount of air poluting agents emitted by cars registered in the region	Federal ministry
313. Доля водохозяйственных участков, класс качества которых (по индексу загрязнения вод) повысился, в общем количестве водохозяйственных участков, расположенных на территории субъекта Российской Федерации	The share of water sites where improvement in water pollution class was registered in the current year in the total area of water sites	Federal ministry
314. Доля рекультивированных земель в общей площади земель, подвергшихся нарушению, включая земли, подвергшиеся радиоактивному и химическому загрязнению	The share of recultivated land sites in the total area of polluted land sites in the region (including radeoactive and chemical pollution)	Regional administration
315. Доля использованных, обезвреженных отходов в общем объеме образовавшихся отходов в процессе производства и потребления	The share of recicled waste in the total amount of waste generated in the region	Federal ministry
316. Доля площади территории субъекта Российской Федерации, занятой особо охраняемыми природными территориями, в общей площади территории субъекта Российской Федерации	The share of the land of the region with protected natural habitat in the total area of the region	Federal ministry
317. Расходы консолидированного бюджета субъекта Российской Федерации на охрану окружающей среды	Budget expenditure on preserving natural habitat	Federal ministry
318. Расходы консолидированного бюджета субъекта Российской Федерации на охрану окружающей среды в части расходов на реализацию региональных программ в области охраны окружающей среды	The share of budget expenditure on preserving natural habitats realised via government programmes in the total amount of budget expenditure on preserving natural habitat	Regional administration
319. Общий объем средств, поступивших в бюджет субъекта Российской Федерации в виде платы за негативное воздействие на окружающую среду, денежных взысканий (штрафов) за нарушение законодательства в области охраны окружающей среды, сумм по искам о возмещении вреда, причиненного окружающей среде	The total amount of revenue from fees, charges and penalties for polution	Regional administration

Russian title	English title	Affiliation
32. Протяженность электрических сетей, в целях увеличения их пропускной способности, введенных в эксплуатацию в соответствии с утвержденными инвестиционными программами	The length of electrical network brought into service in accordance with approved investment programmes	Regional administration
33. Протяженность планируемых к вводу в эксплуатацию в соответствии с утвержденными инвестиционными программами электрических сетей, в целях увеличения их пропускной способности	The length of electrical network planned to be brought into service in accordance with approved investment programmes	Regional administration
34. Наличие в субъекте Российской Федерации утвержденных схем (схемы) территориального планирования субъекта Российской Федерации	Availability of an approved scheme of territorial development in the region	Regional administration
35. Доля городских округов и городских поселений с численностью населения более 50 тыс. человек, в которых приняты генеральные планы (внесены в них изменения) с 1 января 2005 г., в общем количестве городских округов и городских поселений с численностью населения более 50 тыс. человек	The share of municipalities with over 50 000 people which have approved a general devepment plan in the total number of municipalities	Regional administration
36. Доля городских округов и городских поселений с численностью населения более 50 тыс. человек, в которых приняты правила землепользования и застройки, в общем количестве городских округов и городских поселений с численностью населения более 50 тыс. человек	The share of municipalities with over 50 000 people which have implemented general rules of development and land usage in the total number of municipalities	Regional administration
37. Средняя продолжительность периода с даты подачи заявки на предоставление земельного участка в аренду для строительства (кроме жилищного) до даты принятия решения о предоставлении земельного участка в аренду для строительства (кроме жилищного)	The average time between the date of application for development (excluding housing) and the date of making a decision allowing development	Regional administration
38. Средняя продолжительность периода с даты принятия решения о предоставлении земельного участка в аренду для строительства (кроме жилищного) до даты выдачи разрешения на строительство (кроме жилищного)	The average time between the date of making a decision allowing development and the date of issuing a construction permit (excluding housing)	Regional administration
39. Средняя продолжительность периода с даты выдачи разрешения на строительство (кроме жилищного) до даты получения разрешения на ввод объекта капитального строительства в эксплуатацию	The average time between the date of issuing a construction permit and the date of bringing the object in operation (excluding housing)	Regional administration

Russian title	English title	Affiliation
40. Доля земельных участков в субъекте Российской Федерации, предоставленных для строительства (кроме жилищного) по результатам торгов, в общей площади земельных участков в субъекте Российской Федерации, предоставленных для строительства (кроме жилищного)	The share of land sites granted for development through tendering in the total number of land sites grantef for development	Regional administration
40.1. Доля земельных участков, находящихся в государственной собственности субъектов Российской Федерации, муниципальной собственности, а также государственная собственность на которые не разграничена, право постоянного (бессрочного) пользования которыми переоформлено в соответствии с требованиями Федерального закона "О введении в действие Земельного кодекса Российской Федерации" в общем количестве земельных участков, находящихся в государственной собственности субъектов Российской Федерации, муниципальной собственности, а также государственная собственность на которые не разграничена, право постоянного (бессрочного) пользования на которые подлежит переоформлению	The share of land sites with undecided ownerhip rights in the total number of land sites	Regional administration
41. Количество органов исполнительной власти, предприятий и организаций, чье согласование необходимо получить в период предоставления земельного участка в аренду для строительства (кроме жилищного) начиная с даты подачи заявки на предоставление земельного участка в аренду для строительства (кроме жилищного) до даты выдачи разрешения на ввод объекта капитального строительства в эксплуатацию	The number of bureaus whose permission is required to begin development	Regional administration
42. Доля городских округов и городских поселений с численностью населения более 50 тыс. человек, в которых утверждены программы комплексного развития систем коммунальной инфраструктуры, в общем количестве городских округов и городских поселений с численностью населения более 50 тыс. человек	The share of urban districts with over 50 000 population that have adopted comprehensive development programmes in the total number of urban districts with over 50 000 population	Regional administration
45. Младенческая смертность, число умерших в возрасте до 1 года на 1 тыс. родившихся живыми	Infant mortality	Rosstat (Fed.stat.bureau)
46. Смертность населения в возрастной группе от 1 года до 4 лет	Mortality rate (1-4 y.o.)	Rosstat (Fed.stat.bureau)

Russian title	English title	Affiliation
47. Смертность населения в возрастной группе от 5 до 9 лет	Mortality rate (5-9 y.o.)	Rosstat (Fed.stat.bureau)
48. Смертность населения в возрастной группе от 10 до 14 лет	Mortality rate (10-19 y.o.)	Rosstat (Fed.stat.bureau)
49. Смертность населения в возрастной группе от 15 до 19 лет	Mortality rate (15-19 y.o.)	Rosstat (Fed.stat.bureau)
5. Уровень безработицы (по методологии Международной организации труда) в среднем за год	Unemployement rate (average annual)	Rosstat (Fed.stat.bureau)
5.1. Коэффициент напряженности на рынке труда	Labor market tension coefficient	Regional administration
50. Материнская смертность	Maternal mortality	Rosstat (Fed.stat.bureau)
51. Смертность населения трудоспособного возраста	Working population mortality rate	Rosstat (Fed.stat.bureau)
52. Смертность населения трудоспособного возраста от внешних причин	Working population mortality rate (external causes)	Rosstat (Fed.stat.bureau)
53. Смертность населения трудоспособного возраста от болезней системы кровообращения	Working population mortality rate (blood circulation)	Rosstat (Fed.stat.bureau)
54. Смертность населения трудоспособного возраста от новообразований	Working population mortality rate (cancer)	Rosstat (Fed.stat.bureau)
55. Смертность населения в результате дорожно-транспортных происшествий	Working population mortality rate (traffic accidents)	Rosstat (Fed.stat.bureau)
56. Общая численность лиц, впервые признанных инвалидами	Total number of people registered as disabled	Federal ministry
57. Число лиц трудоспособного возраста, впервые признанных инвалидами	The number of working age people registered as disabled for the first time	Federal ministry
58. Удовлетворенность населения медицинской помощью	Citizen satisfaction with medical services	Federal security service
59. Количество обоснованных жалоб на отказ в оказании медицинской помощи, предоставляемой в рамках территориальной программы обязательного медицинского страхования	The number of complaints regarding refusal to provide medical services provided by territorial programme of compulsory medical ensurance	Federal ministry
59.1. Доля граждан, получивших обоснованный отказ в оказании высокотехнологичной медицинской помощи, в общей численности граждан, направленных на оказание высокотехнологичной медицинской помощи органом управления здравоохранением субъекта Российской Федерации	The share of patients who was refused high-tech medical services in the total number of patients referred to specialised health service organisations of the region	Federal ministry
59.2. Заболеваемость населения туберкулезом	Tuberculosis sickness rate	Federal ministry
59.3. Смертность населения от туберкулеза	Tuberculosis mortality rate	Federal ministry

Russian title	English title	Affiliation
6. Доля трудоустроенных граждан в общей численности граждан, обратившихся за содействием в государственные учреждения занятости с целью поиска подходящей работы	The share of applicants who found employment in the total number of applicants of social security services.	Regional administration
61. Объем оказанной стационарной медицинской помощи в расчете на 1 жителя	The volume of in-patient medical services provided per patient	Federal ministry
62. Объем оказанной амбулаторной медицинской помощи в расчете на 1 жителя	The volume of out-patient medical services provided per patient	Federal ministry
63. Объем оказанной скорой медицинской помощи в расчете на 1 жителя	The volume of medical services provided per patient	Federal ministry
64. Объем оказанной медицинской помощи в дневных стационарах всех типов в расчете на 1 жителя	The volume of medical services provided in hospitals of all types per patient	Federal ministry
65. Стоимость единицы объема оказанной стационарной медицинской помощи (фактическое значение)	Unit cost of in-patient health services (factual)	Federal ministry
66. Стоимость единицы объема оказанной амбулаторной медицинской помощи (фактическое значение)	Unit cost of out-patient health services (factual)	Federal ministry
67. Стоимость единицы объема оказанной медицинской помощи в дневных стационарах всех типов (фактическое значение)	Unit cost of daily in-patient health services (factual)	Federal ministry
68. Стоимость единицы объема оказанной скорой медицинской помощи (фактическое значение)	Unit cost of health services (factual)	Federal ministry
69.1. Количество государственных (муниципальных) учреждений социального обслуживания	The number of state (municipal) social security service providers	Federal ministry
7. Доля трудоустроенных граждан, относящихся к категории инвалидов, в общей численности граждан, относящихся к категории инвалидов, обратившихся за содействием в государственные учреждения занятости с целью поиска подходящей работы	The share of disabled people who found employment in the total number of disabled people who applied for unemployment services	Regional administration
70. Количество государственных (муниципальных) учреждений здравоохранения, здания которых находятся в аварийном состоянии или требуют капитального ремонта	The number of state (municipal) health service institutions whose buildings require capital renovation	Federal ministry
70.1. Количество государственных (муниципальных) учреждений социального обслуживания, здания которых находятся в аварийном	The number of state (municipal) social security service institutions whose buildings require capital	Federal ministry

Russian title	English title	Affiliation
состоянии или требуют капитального ремонта	renovation	
71. Среднемесячная номинальная начисленная заработная плата врачей государственных (муниципальных) учреждений здравоохранения	Average monthly nominal salary of doctors at state (municipal) health service institutions	Regional administration
72. Среднемесячная номинальная начисленная заработная плата среднего медицинского персонала государственных (муниципальных) учреждений здравоохранения	Average monthly nominal salary of middle-tier staff at state (municipal) health service institutions	Regional administration
72.1. Среднемесячная номинальная начисленная заработная плата прочего персонала государственных (муниципальных) учреждений здравоохранения	Average monthly nominal salary of other staff at state (municipal) health service institutions	Regional administration
72.2. Среднемесячная номинальная начисленная заработная плата прочего персонала, в том числе младшего медицинского персонала государственных (муниципальных) учреждений здравоохранения	Average monthly nominal salary of other staff including junior staff at state (municipal) health service institutions	Regional administration
73. Численность работающих в государственных (муниципальных) учреждениях здравоохранения: физические лица (за исключением работающих в федеральных учреждениях)	The number of staff at state (municipal) health service institutions (excluding federal institutions)	Federal ministry
73.1. Численность прочего персонала (физических лиц) в государственных (муниципальных) учреждениях здравоохранения (за исключением лиц, работающих в федеральных учреждениях)	The number of other staff at state (municipal) health service insitutions	Federal ministry
73.2. Численность младшего медицинского персонала в государственных (муниципальных) учреждениях здравоохранения	The number of junior medical staff at state (municipal) health service institutions	Federal ministry
74. Численность врачей (физические лица) в государственных (муниципальных) учреждениях здравоохранения	The number of doctors at state (municipal) health service institutions	Federal ministry
75. Численность среднего медицинского персонала (физические лица) в государственных (муниципальных) учреждениях здравоохранения	The number of middle-tier medical staff at state (municipal) health service institutions	Federal ministry
76. Количество коек в государственных (муниципальных) учреждениях здравоохранения в городской местности	The number of beds at state (municipal) health service institutions in urban areas	Federal ministry
76.1. Количество коек в государственных (муниципальных) учреждениях здравоохранения	The number of beds at state (municipal) health service institutions	Federal ministry

Russian title	English title	Affiliation
77. Количество коек в государственных (муниципальных) учреждениях здравоохранения в сельской местности	The number of beds at state (municipal) health service institutions in rural areas	Federal ministry
78. Средняя продолжительность 1 случая временной нетрудоспособности в связи с заболеванием	Average length of one case of temporary disability	Federal ministry
79. Средняя продолжительность пребывания пациента на койке в государственных (муниципальных) учреждениях здравоохранения	Average length of stay of one patient at state (municipal) healt care institutions	Federal ministry
8. Доля населения с денежными доходами ниже региональной величины прожиточного минимума в общей численности населения субъекта Российской Федерации	The share of population with income below the poverty line	Rosstat (Fed.stat.bureau)
80. Среднегодовая занятость койки в государственных (муниципальных) учреждениях здравоохранения	Average annual occupation rate of hospital beds	Federal ministry
81. Уровень госпитализации в государственные (муниципальные) учреждения здравоохранения	The level of admittance to state (municipal) hospitals	Federal ministry
82. Стоимость 1 койко-дня в государственных (муниципальных) учреждениях здравоохранения (без учета видов расходов, возмещаемых в рамках территориальной программы обязательного медицинского страхования)	The cost of one bed-day at state (municipal) health care establishments	Federal ministry
82.1. Фактическая стоимость 1 койко-дня в государственных (муниципальных) учреждениях здравоохранения без учета расходов на оплату труда и начислений на оплату труда	Factual cost of one bed-day at state (municipal) hospitals (excluding salaries)	Federal ministry
82.2. Фактическая стоимость вызова скорой медицинской помощи без учета расходов на оплату труда и начислений на оплату труда	Factual cost of one emergence call (ambulance) excluding salaries	Federal ministry
83. Расходы консолидированного бюджета субъекта Российской Федерации на здравоохранение: всего	Budget expenditure on health services: total	Federal ministry
84. Расходы консолидированного бюджета субъекта Российской Федерации на здравоохранение в части увеличения стоимости основных средств	Budget expenditure on health services: capital expenditure	Federal ministry
85. Расходы консолидированного бюджета субъекта Российской Федерации на здравоохранение в части текущих расходов	Budget expenditure on health services: current expenditure	Federal ministry

Russian title	English title	Affiliation
86. Расходы консолидированного бюджета субъекта Российской Федерации на здравоохранение в части текущих расходов на оплату труда и начислений на оплату труда	Budget expenditure on health services: salaries	Federal ministry
87. Расходы консолидированного бюджета субъекта Российской Федерации на реализацию территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации в расчете на 1 жителя	Budget expenditure on health service through the territorial programme of state warrants of free medical services per capita	Federal ministry
88. Расходы средств обязательного медицинского страхования в расчете на 1 жителя	Expenditure of compulsory medical insurance funds per capita	Federal ministry
88.1. Расходы территориальных государственных внебюджетных фондов	Expenditure of territorial state off-budget funds	Federal ministry
89. Объем расходов на оказание скорой медицинской помощи в рамках территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации	Costs of emergency health services according to the territorial programme of free medical services	Federal ministry
9. Реальная среднемесячная начисленная заработная плата работников	Real monthly accrued salary of workers	Rosstat (Fed.stat.bureau)
90. Объем расходов на оказание амбулаторной медицинской помощи в рамках территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации	Costs of out-patient care according to the territorial programme of free medical services	Federal ministry
91. Объем расходов на оказание стационарной медицинской помощи в рамках территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации	Costs of in-patient care according to the territorial programme of free medical services	Federal ministry
92. Объем расходов на оказание медицинской помощи в дневных стационарах всех типов в рамках территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации	Costs of hospital care according to the territorial programme of free medical services	Federal ministry
93. Фактическая стоимость территориальной программы обязательного медицинского страхования	Factual cost of the territorial programme of compulsory medical insurance	Federal ministry
93.1. Дефицит финансового обеспечения территориальной программы государственных гарантий оказания бесплатной медицинской помощи гражданам Российской Федерации	Deficit of the programme of state warrants of medical service provision	Federal ministry
94. Количество государственных и муниципальных учреждений	The number of state and municipal health service	Federal ministry

Russian title	English title	Affiliation
здравоохранения: всего	institutions	
94.1. Количество государственных (муниципальных) учреждений здравоохранения особого типа	The number of state and municipal specialized health service institutions	Federal ministry
95. Количество государственных и муниципальных амбулаторно-поликлинических учреждений	The number of state and municipal outpatients' clinic	Federal ministry
96. Количество государственных и муниципальных больничных учреждений	The number of state and municipal hospitals	Federal ministry
97. Количество государственных (муниципальных) учреждений здравоохранения, переведенных преимущественно на одноканальное финансирование (не менее 70% от общего объема финансирования за счет средств обязательного медицинского страхования) через систему обязательного медицинского страхования	The number of health service institutions where one-channel financing has been implemented (over 70% funded through compulsory medical insurance)	Federal ministry
98. Количество государственных (муниципальных) учреждений здравоохранения, использующих единые информационные технологии учета объемов и стоимости оказанной медицинской помощи	The number of state (municipal) health service institutions using unified information systems quantifying amount and costs of health services provided	Regional administration
99. Количество государственных (муниципальных) учреждений здравоохранения, переведенных на новую (отраслевую) систему оплаты труда, ориентированную на результат	The number of health service institutions where new performance pay sheme have been implemented	Federal ministry

Appendix 7. Indicators included in the survey of municipalities

N	Пара из опросника		Clustering index [54]	Trust index[55]
1	Среднемесячная номинальная начисленная заработная плата работников муниципальных учреждений физической культуры и спорта	average nominal salary of staff of sport establishments	0,17	1,51
2	Расходы бюджета муниципального образования на общее образование в расчете на 1 обучающегося в муниципальных общеобразовательных учреждениях	average cost of one class in high schools in the region	0,21	1,61
3	Доля среднесписочной численности работников (без внешних совместителей) малых и средних предприятий в среднесписочной численности работников (без внешних совместителей) всех предприятий и организаций	the share employees of small and medium sided businesess in the total number of emplyees	0,22	0,74
4	Среднемесячная номинальная начисленная заработная плата учителей муниципальных общеобразовательных учреждений	average nominal salary of teachers in state owned high schools	0,24	1,61
5	Среднемесячная номинальная начисленная заработная плата работников муниципальных общеобразовательных учреждений	average nominal monthly salary of staff of state owned high schools	0,27	1,58
6	Объем не завершенного в установленные сроки строительства, осуществляемого за счет средств бюджета городского округа (муниципального района)	the share of unfinished construction financed by the regional budget	0,27	1,31
7	Среднемесячная номинальная начисленная заработная плата работников муниципальных учреждений культуры и искусства	average nominal monthly salary of staff of cultural establishments	0,27	1,57
8	Доля прибыльных сельскохозяйственных организаций в общем их числе	the share of profitable large and medium sized businesses	0,27	1,12
9	Доля выпускников муниципальных общеобразовательных учреждений, сдавших единый государственный экзамен по русскому языку и	the number of high school graduates that passed the national exam in Russian language	0,31	1,61

[54] The index measuring the share of observations of growth indices falling in the interval from -0,3% to +0,3%

[55] The index measuring trust in data reported by other municipalities of the region. Constructed from the responses to Question 5 of the survey of municipalities by assigning the following points to responses Strongly agree +2, agree +1, disagree -1, strongly disagree 2, do not know – 0. Index is the average.

	математике, в общей численности выпускников муниципальных общеобразовательных учреждений, сдававших единый государственный экзамен по данным предметам			
10	Доля налоговых и неналоговых доходов местного бюджета (за исключением поступлений налоговых доходов по дополнительным нормативам отчислений) в общем объеме собственных доходов бюджета муниципального образования (без учета субвенций)	the share of tax and non-tax incomes in the total amount of income of municipal budgets (excluding subventions)	0,31	1,58
11	Доля населения, систематически занимающегося физической культурой и спортом	the number of people systematically doing sports	0,31	1,21
13	Расходы бюджета муниципального образования на содержание работников органов местного самоуправления в расчете на одного жителя муниципального образования	expenditure of the consolidated regional budget (staff payroll)	0,33	1,59
14	Объем инвестиций в основной капитал (за исключением бюджетных средств) в расчете на 1 жителя	capital investment per capita	0,34	1,05
15	Объём расходов бюджета муниципального образования на компенсацию разницы между экономически обоснованными тарифами на жилищно-коммунальные услуги и тарифами, установленными для населения	budget expenditure on communal services - subsidies of tariffs	0,35	1,27
16	Общая площадь жилых помещений (в расчёте на одного жителя), введенная в действие за отчётный год	total area of housing built annually per capita	0,37	1,27
17	Доля расходов бюджета городского округа (муниципального района), формируемых в рамках программ, в общем объеме расходов бюджета городского округа (муниципального района), без учета субвенций на исполнение делегируемых полномочий	the share of expenditure done through public sector programmes in the total expenditure of the regional budget	0,38	1,56
18	Доля протяженности автомобильных дорог общего пользования местного значения, не отвечающих нормативным требованиям, в общей протяженности автомобильных дорог общего пользования местного значения	the share of local low quality roads in the total milege of local roads	0,39	0,99
19	Доля выпускников муниципальных общеобразовательных учреждений, не получивших аттестат о среднем (полном) образовании, в общей численности выпускников муниципальных общеобразовательных учреждений	the number of high school graduates who did not receive the education certificate	0,39	1,62
	Доля обрабатываемой пашни в общей площади пашни муниципального	the share of tilled ploughed field in total land area	0,41	0,90

20	района	of the district		
21	Площадь земельных участков, предоставленных для строительства, в отношении которых с даты принятия решения о предоставлении земельного участка или подписания протокола о результатах торгов (конкурсов, аукционов) не было получено разрешение на ввод в эксплуатацию в течение 3 лет (для объектов жилищного строительства)	the area of land sites for which the time between issuing the development license and completion of construction exeeded 3 years	0,45	1,23
22	Уровень фактической обеспеченности учреждениями физической культуры и спорта в городском округе (муниципальном районе) от нормативной потребности: плоскостными спортивными сооружениями	availability of sport facilities in the region: playing fields	0,45	1,42
23	Доля населения, участвующего в платных культурно-досуговых мероприятиях, организованных органами местного самоуправления городских округов и муниципальных районов	the shar of population taking part in organised cultural events	0,50	1,20
24	Доля многоквартирных домов, расположенных на земельных участках, в отношении которых осуществлен государственный кадастровый учет	the share of appartment blocks located on land sites that have been included in the cadastre	0,50	1,22
25	Уровень фактической обеспеченности учреждениями физической культуры и спорта в городском округе (муниципальном районе) от нормативной потребности: спортивными залами	availability of sport facilities in the region: sport halls	0,52	1,47
26	Удельная величина потребления энергетических ресурсов в многоквартирных домах	unit quantity of electrical power consumption in appartment blocks	0,52	0,88
27	Площадь земельных участков, предоставленных для строительства (всего) в расчете на 10 тыс. человек населения	the area of land licenced for housing development per capita	0,52	1,26
28	Площадь земельных участков, предоставленных для жилищного строительства, индивидуального строительства и комплексного освоения в целях жилищного строительства	the area of land licenced for housing development per capita	0,52	1,35
29	Количество муниципальных услуг, предоставляемых органами местного самоуправления, муниципальными учреждениями в электронном виде	the number of public services provided online	0,56	1,51
30	Доля детей первой и второй групп здоровья в общей численности обучающихся в муниципальных общеобразовательных учреждениях	the share of kids qualified in 1 and 2 health categories in the total number of pupils	0,56	1,40
31	Общая площадь жилых помещений, приходящаяся в среднем на одного жителя, всего	housing area per capita	0,57	1,28
	Доля организаций коммунального комплекса, осуществляющих	the share of communal service providers with less	0,62	1,28

33	производство товаров, оказание услуг по водо-, тепло-, газо-, электроснабжению, водоотведению, очистке сточных вод, утилизации (захоронению) ТБО и использующих объекты коммунальной инфраструктуры на праве частной собственности, по договору аренды или концессии, участие субъекта РФ и (или) городского округа (муниципального района) в уставном капитале которых составляет не более 25 процентов, в общем числе организаций коммунального комплекса, осуществляющих свою деятельность на территории ГО (МР)	than 25% state ownership in the total number of communal service providers		
34	Доля муниципальных общеобразовательных учреждений, здания которых находятся в аварийном состоянии или требуют капитального ремонта, в общем количестве муниципальных общеобразовательных учреждений	the number of high schools requireing renovation	0,80	1,35
35	Число субъектов малого и среднего предпринимательства в расчете на 10 тыс. человек населения	the number of small and medium sized business per 1000 people	0,88	1,00
36	Доля населения, проживающего в многоквартирных домах, признанных в установленном порядке аварийными	the share of population living in houses requiring capital renovation	0,92	1,16
37	Наличие в городском округе (муниципальном районе) утвержденного генерального плана городского округа (схемы территориального планирования муниципального района)	availability of an approved scheme of territorial development in the region	0,98	1,58

Appendix 8. Quotes in original Russian

BIBLIOGRAPHY

Agarev, A.V. (2005), *Tragicheskaya avantura: Sel'skoe khozyaistvo v Ryazanskoy oblasty 1950-1960 (Tragic venture: Agriculture in Ryazan oblast' 1950-1960)* (Ryazan: Russkoe slovo (Ryazobltipografiya)) 192.

Ahn, T. K., Ostrom, E., and Walker, J. M. (2003), 'Heterogeneous preferences and collective action', *Public Choice,* 117 (3-4), 295-314.

Andrews, R., et al. (2008), 'Organizational strategy, external regulation and public service performance', *Public Administration,* 86 (1), 185-203.

Andrews, R., et al. (2009), 'Centralization, Organizational Strategy, and Public Service Performance', *Journal of Public Administration Research and Theory,* 19 (1), 57-80.

Audit Commission (1999), *A measure of success : setting and monitoring local performance targets* (London: Audit Commission for Local Authorities and the National Health Service in England and Wales).

Barabashev, Alexei and Straussman, Jeffrey D. (2007), 'Public Service Reform in Russia, 1991-2006', *Public Administration Review,* 67 (3), 373-82.

Barabashev, Alexei and Klimenko, A.V. (2010), 'Administrativnaya reforma i reforma gosudarstvennoy sluzhby v Rossii - voprosy realizacii i koordinacii (Administrative reform and civil service reform in Russia - issues of realisation and coordination)', *Gosudarstvennoe i municipalnoye upravlenie, Preprint WP8/2010/01* (Moscow: National research university - Higher school of economics).

Barber, Michael (2007), *Instruction to deliver : Tony Blair, public services and the challenge of achieving targets / Michael Barber* (London: Politico's) xiii, 418.

Bardach, E. (2001), 'Developmental Dynamics: Interagency Collaboration as an Emergent Phenomenon', *Journal of Public Administration Research & Theory,* 11 (2), 149.

Barrett, S. M. (2004), 'Implementation studies: Time for a revival? Personal reflections on 20 years of implementation studies', *Public Administration,* 82 (2), 249-62.

Batley, R. and Larbi, G. (2004), *Changing Role of Government : the Reform of Public Services in Developing Countries* (Palgrave).

Beatty, A. L., Ke, B., and Petroni, K. R. (2002), 'Earnings management to avoid earnings declines across publicly and privately held banks', *Accounting Review,* 77, 547-70.

Behn, R. D. (2003), 'Why measure performance? Different purposes require different measures', *Public Administration Review,* 63 (5), 586-606.

Bennett, B. K. and Bradbury, M. E. (2010), 'An analysis of the reasons for the asymmetries surrounding earnings benchmarks', *Accounting and Finance,* 50, 529-54.

Bevan, G. and Hood, C. (2006), 'What's measured is what matters: Targets and gaming in the English public health care system', *Public Administration,* 84 (3), 517-38.

Bird, S. M., et al. (2005), 'Performance indicators: Good, bad, and ugly', *Journal of the Royal Statistical Society Series a-Statistics in Society,* 168, 1-27.

Blokhin, A.A. (2011), 'Filosofiya rezultativnosti v reformakh gosudarstvennogo upravleniya v Rossii (The philosophy of management by results in publis sector reforms in Russia)', *Strategicheskoe upravlenie: ot idei do rezultata* (Moscow: Analiticheskiy centr pri Pravitelstve Rossiiskoy Federazii), 164.

Bogason, Peter and Zølner, Mette (2007), *Methods in democratic network governance / edited by Peter Bogason and Mette Zølner* (Basingstoke: Palgrave Macmillan) xiv, 249.

Bohte, J. and Meier, K. J. (2000), 'Goal displacement: Assessing the motivation for organizational cheating', *Public Administration Review,* 60 (2), 173-82.

Bouckaert, Geert and Halachmi, A. (1994), 'Performance Measurement, Organizational Technology and Organizational Design', *Work Study,* 43 (3), 19-25.

Bouckaert, Geert and Peters, B. Guy (2002), 'Performance Measurement and Management: The Achilles' Heel in Administrative Modernization', *Public Performance & Management Review,* 25 (4), 359-62.

Boyne, G. A. (1998), 'Bureaucratic theory meets reality: Public choice and service contracting in US local government', *Public Administration Review,* 58 (6), 474-84.
Boyne, G. A. and Chen, A. A. (2007), 'Performance targets and public service improvement', *Journal of Public Administration Research and Theory,* 17 (3), 455-77.
Boyne, G. A., et al. (2002), 'Plans, performance information and accountability: The case of best value', *Public Administration,* 80 (4), 691-710.
Boyne, G. A., et al. (2005), 'Where next? Research directions on performance in public organizations', *Journal of Public Administration Research and Theory,* 15 (4), 633-39.
Breul, Jonathan D. (2007), 'GPRA—A foundation for performance budgeting', *Public Performance and Management Review,* 30 (3), 312-31.
Brunsson, Nils (1989), *The organization of hypocrisy : talk, decisions and actions in organizations / Nils Brunsson*

translated by Nancy Adler (Chichester: Chichester : Wiley).
Bryman, Alan (2006a), 'Paradigm Peace and the Implications for Quality', *International Journal of Social Research Methodology,* 9 (2), 111-26.
--- (2006b), 'Integrating quantitative and qualitative research: how is it done?', *Qualitative Research,* 6 (1), 97-113.
Buckley, Noah, Garifullina, Guzel, and Reuter, Ora John (2012), 'Bureaucratic appointments in hybrid regimes', *Higher School of Economics Research Paper,* WP BRP 11/PS/2012.
Burgstahler, D. and Dichev, I. (1997), 'Earnings management to avoid earnings decreases and losses', *Journal of Accounting & Economics,* 24, 99-126.
Burrell, G. and Morgan, Gareth (1979), *Sociological Paradigms and Organizational Analysis* (London: Heinemann).
Carter, N. (1991), 'LEARNING TO MEASURE PERFORMANCE - THE USE OF INDICATORS IN ORGANIZATIONS', *Public Administration,* 69 (1), 85-101.
Charbonneau, E. and Bellavance, F. (2012), 'Blame Avoidance In Public Reporting: Evidence from a Provincially Mandated Municipal Performance Measurement Regime', *Public Performance & Management Review,* 35 (3), 399-421.
Chekharin, E. M. (1977), *The Soviet political system under developed Socialism* (Progress socialism today; Moscow: Progress Publishers) 346 p.
Clarkson, P., et al. (2009), 'Has social care performance in England improved? An analysis of performance ratings across social services organisations', *Policy Studies,* 30 (4), 403-22.
Committee of civic initiatives (Komitet grazhdanskikh iniziativ) (2012), 'Pravookhranitel'naya deyatel'nost' v Rossii: struktura, funkcionirovanie, puti reformirovaniya (Policing in Russia: structure, functions and directions of reforms)'.
Cunliffe, Ann L. (2011), 'Crafting Qualitative Research: Morgan and Smircich 30 Years On', *Organizational Research Methods,* 14 (4), 647-73.
Darden, Keith (2008), 'The Integrity of Corrupt States: Graft as an Informal State Institution', *Politics & Society,* 36 (1), 35-59.
Davies, M. (2004), 'Performance measurement in the UK public sector: Understanding performance indicators.', *Journal of Finance and Management in Public Services,* 3 (2), 31-47.
de Bruijn, Hans (2002), 'Performance measurement in the public sector: Strategies to cope with the risks of performance measurement', *The International Journal of Public Sector Management,* 15 (6/7), 578-94.
Dechow, P. M., Sloan, R. G., and Sweeney, A. P. (1995), 'DETECTING EARNINGS MANAGEMENT', *Accounting Review,* 70, 193-225.
Degeorge, F., Patel, J., and Zeckhauser, R. (1999), 'Earnings management to exceed thresholds', *Journal of Business,* 72, 1-33.
Dietz, T., Ostrom, E., and Stern, P. C. (2003), 'The struggle to govern the commons', *Science,* 302 (5652), 1907-12.
Downs, Anthony (1967), *Inside bureaucracy* (Boston: Little Brown and Co.) xv, 292.

Dubnick, Melvin (2005), 'Accountability and the Promise of Performance: In Search of the Mechanisms', *Public Performance and Management Review,* 28 (3), 376-417.
Dull, M. (2006), 'Why PART? The institutional politics of presidential budget reform', *Journal of Public Administration Research and Theory,* 16 (2), 187-215.
Durtschi, Cindy, Hillison, William, and Pacini, Carl (2004), 'The effective use of Benford's law to assist in detecting fraud in accounting data', *Journal of forensic accounting,* 5 (1), 17-34.
Eames, M. and Kim, Y. (2012), 'Analyst vs. Market Forecasts of Earnings Management to Avoid Small Losses', *Journal of Business Finance & Accounting,* 39, 649-74.
Eisenhardt, Kathleen M. (1989), 'Building Theories from Case Study Research', *The Academy of Management Review,* 14 (4), 532-50.
Elmore, R. F. (1980), 'BACKWARD MAPPING - IMPLEMENTATION RESEARCH AND POLICY DECISIONS', *Political Science Quarterly,* 94 (4), 601-16.
Enikolopov, Ruben, et al. (2012), 'Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections', *Proceedings of the National Academy of Sciences*.
Frank, H. M. Verbeeten (2008), 'Performance management practices in public sector organizations', *Accounting, Auditing & Accountability Journal,* 21 (3), 427-54.
Frye, Timothy , Reuter, Ora John, and Buckley, Noah (2011), 'The Political Economy of Russian Gubernatorial Election and Appointment', *Higher School of Economics Research Paper*.
Game, Chris (2006), 'Comprehensive performance assessment in English local government', *International Journal of Productivity and Performance Management,* 55 (6), 466-79.
Gilmour, J. B. (2008), 'Implementing OMB's Program Assessment Rating Tool', in T. Buss F. Shea (ed.), *Performance management and budgeting: How governments can learn from experience* (Armonk, NY: M.E. Sharpe), 21–48.
Gilmour, J. B. and Lewis, D. E. (2006), 'Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores', *Public Administration Review,* 66 (5), 742-52.
Gimpelson, V., Magun, V., and Brym, R. J. (2009), 'Hiring and Promoting Young Civil Servants: Weberian Ideals versus Russian Reality', in Don Karl Rowney and Eugene Huskey (eds.), *Russian bureaucracy and the state: officialdom from Alexander III to Vladimir Putin* (Basingstoke: Palgrave Macmillan).
Gioia, Dennis A. and Pitre, Evelyn (1990), 'Multiparadigm Perspectives on Theory Building', *The Academy of Management Review,* 15 (4), 584-602.
Goffman, Erving (2006), 'On face-work: an analysis of ritual elements in social interaction', in Adam Jaworski and Nikolas Coupland (eds.), *The discourse reader* (2nd edn.; London ; New York: Routledge), xvi, 560 p.
Goldfinch, Shaun and Wallis, J. O. E. (2010), 'TWO MYTHS OF CONVERGENCE IN PUBLIC MANAGEMENT REFORM', *Public Administration - Oxford,* 88 (4), 1099-115.
Gray, Andrew and Jenkins, Bill (1995), 'From public administration to public management: Reassessing a revolution?', *Public Administration,* 73 (1), 75-99.
Greitens, Thomas J. and Joaquin, M. Ernita (2010), 'POLICY TYPOLOGY AND PERFORMANCE MEASUREMENT Results from the Program Assessment Rating Tool (PART)', *Public Performance and Management Review,* 33 (4), 555-70.
Gueorguieva, V., et al. (2009), 'The Program Assessment Rating Tool and the Government Performance and Results Act Evaluating Conflicts and Disconnections', *American Review of Public Administration,* 39 (3), 225-45.
Halachmi, A. (2002a), 'Performance Measurement, Accountability, and Improved Performance', *Public Performance & Management Review,* 25 (4), 370-74.
--- (2002b), 'Performance measurement: a look at some possible dysfunctions', *Work Study,* 51 (5), 230-39.
Hardin, Garrett (1968), 'The Tragedy of the Commons', *Science,* 162 (3859), 1243-48.
Heinrich, Carolyn J. (2002), 'Outcomes-Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness', *Public Administration Review,* 62 (6), 712-25.

Hood, C. (1991), 'A public management for all seasons?', *Public Administration,* 69 (1), 3-19.
--- (2002), 'The risk game and the blame game', *Government and Opposition,* 37 (1), 15-37.
--- (2006), 'Gaming in targetworld: The targets approach to managing British public services', *Public Administration Review,* 66 (4), 515-21.
--- (2007a), 'What happens when transparency meets blame-avoidance?', *Public Management Review,* 9 (2), 191-210.
--- (2007b), 'Public service management by numbers: Why does it vary? Where has it come from? What are the gaps and the puzzles?', *Public Money & Management,* 27 (2), 95-102.
--- (2011), *Blame Game: Spin, Bureaucracy, and Self-Preservation in Government* (Princeton University Press) 240.
Hood, C. and Jackson, Michael (1994), 'Keys for Locks in Administrative Argument', *Administration & Society,* 25 (4), 467-88.
Hood, C. and Peters, Guy (2004), 'The Middle Aging of New Public Management: Into the Age of Paradox?', *Journal of Public Administration Research & Theory,* 14 (3), 267-82.
Hood, C. and Dixon, Ruth (2010), 'The Political Payoff from Performance Target Systems: No-Brainer or No-Gainer?', *Journal of Public Administration Research and Theory,* 20, I281-I98.
Hunt, Michael (1995), 'Openness and Civil Service Ethics', *International Review of Administrative Sciences,* 61 (1), 11-16.
Hutton, William L. and Massey, Andrew (2006), 'Professional Ethics and Public Service: Can Professionals Serve Two Masters?', *Public Money & Management,* 26 (1), 23-30.
Hyndman, Noel and Eden, Ron (2000), 'A study of the coordination of mission, objectives and targets in U.K. executive agencies', *Management Accounting Research,* 11 (2), 175-91.
--- (2001), 'Rational management, performance targets and executive agencies: Views from agency chief executives in northern ireland', *Public Administration,* 79 (3), 579-98.
Imperial, M. (2005), 'Using collaboration as a governance strategy - Lessons from six watershed management programs', *Administration & Society,* 37 (3), 281-320.
Institute of the national project «Obshestvenniy dogovor» (2011), 'Gotovnost' grazhdan prinimat' na sebya obyazatel'stva po pokritiyu izderzhek po proizvodstvu obshestvennikh blag i opredeleniu sostava i kachestva obshestvennikh blag, proizvodimykh gosudarstvom. (The readiness of citizens to take part in the production of public goods or bearing the costs of public goods produced by the government) \', (Moscow: Institute of the national projects "Obshestvenniy dogovor").
Ittner, Christopher D. and Larcker, David F. (2001), 'Assessing empirical research in managerial accounting: a value-based management perspective', *Journal of Accounting and Economics,* 32 (1–3), 349-410.
Jackson, P. M. (2011), 'Governance by numbers: what have we learned over the past 30 years?', *Public Money & Management,* 31 (1), 13-26.
Jakobson, Lev (2001), 'Public management in Russia: changes and inertia', *International Public Management Journal,* 4 (1), 27-48.
Jansen, I. P., Ramnath, S., and Yohn, T. L. (2012), 'A Diagnostic for Earnings Management Using Changes in Asset Turnover and Profit Margin', *Contemporary Accounting Research,* 29, 221-51.
Jones, C., Hesterly, W. S., and Borgatti, S. P. (1997), 'A general theory of network governance: Exchange conditions and social mechanisms', *Academy of Management Review,* 22 (4), 911-45.
Jones, J. J. (1991), 'EARNINGS MANAGEMENT DURING IMPORT RELIEF INVESTIGATIONS', *Journal of Accounting Research,* 29, 193-228.
Jones, K. L., Krishnan, G. V., and Melendrez, K. D. (2008), 'Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical analysis', *Contemporary Accounting Research,* 25, 499-531.
K.E., Yunusova (2009), 'Socialno-ekonomicheskie prozessy i dukhovnaya zhizn' v Uzbekistane v 80-kh godakh XX veka (Social-ekonomic processes and spiritual life in Uzbekistan in 1980s) ', (Akademiya nauk respubliki Uzbekistan).

Kaplan, R. S. (2001), 'Strategic Performance Measurement and Management in Nonprofit Organizations', *Nonprofit Management and Leadership,* 11 (3), 353-70.
Kaplan, R. S. and Norton, D. P. (1992), 'The balanced scorecard - measures that drive performance', *Harvard Business Review,* 70 (1), 71-79.
Kelly, J. M. (2002), 'Why We Should Take Performance Measurement on Faith (Facts Being Hard to Come by and Not Terribly Important)', *Public Performance & Management Review,* 25 (4), 375-80.
Keraudren, Philippe (1995), 'Administrative Reform, Ethics and Openness: The Balance between Effectiveness and Administrative Identity', *International Review of Administrative Sciences,* 61 (1), 41-60.
Kernaghan, Kenneth (2003), 'Integrating Values into Public Service: The Values Statement as Centerpiece', *Public Administration Review,* 63 (6), 711-19.
Khabaev, S. G. (2010), 'Analiz praktiki BOR v zdravookhranenii (Analysing the practice of performance based budgeting in health service)', *budget.ru*. <http://bujet.ru/article/83969.php>.
Kloot, Louise and Martin, John (2000), 'Strategic performance management: A balanced approach to performance management issues in local government', *Management Accounting Research,* 11 (2), 231-51.
Koch-Weser, Iacob N. (2013), 'The Reliability of China's Economic Data: An Analysis of National Output', *U.S.-China Economic and Security Review Commission Staff Research Project*.
Kravchuk, Robert S. and Schack, Ronald W. (1996), 'Designing effective performance-measurement systems under the Government Performance and Results', *Public Administration Review,* 56 (4), 348.
Kuzminov, Y. and Yudkevich, M. (2010), 'Za predelami rynka: instituty upravleniya tranzakciyami v slozhnom mire (Beyond markets: managing transactions in a complex world)', *Voprosy Ekonomiki,* 1.
Kvale, S. (1996), *InterViews : an Introduction to Qualitative Research Interviewing* (London: Sage Publications).
Lapsley, Irvine (1999), 'Accounting and the New Public Management: Instruments of substantive efficiency or a', *Financial Accountability & Management,* 15 (3/4), 201.
Lewis, Marianne W. and Grimes, Andrew J. (1999), 'Metatriangulation: Building Theory from Multiple Paradigms', *The Academy of Management Review,* 24 (4), 672-90.
Libman, A., Kozlov, Vladimir, and Schultz, André (2012a), 'Roving Bandits in Action: Outside Option and Governmental Predation in Autocracies', *Kyklos,* 65 (4), 526-62.
Libman, A., Kozlov, V. , and Schultz, A. (2012b), 'Roving Bandits in Action: Outside Option and Governmental Predation in Autocracies', (Frankfurt School - Working paper series №190).
Likierman, Andrew (1993), 'Performance Indicators: 20 Early Lessons from Managerial Use', *Public Money & Management,* 13 (4), 15-22.
Malkov, K.B. (2010), 'Upravlenie po rezultatav v organakh gosudarstvennoy vlasti (Performance management in the executive branch of power)', (MSU Lomonosov).
McGinnis, M. D. and Ostrom, E. (2012), 'Reflections on Vincent Ostrom, Public Administration, and Polycentricity', *Public Administration Review,* 72 (1), 15-25.
Meekings, A. (1995), 'UNLOCKING THE POTENTIAL OF PERFORMANCE-MEASUREMENT - A PRACTICAL IMPLEMENTATION GUIDE', *Public Money & Management,* 15 (4), 5-12.
Meier, K. J. and O'Toole, L. J. (2006), 'Political control versus bureaucratic values: Reframing the debate', *Public Administration Review,* 66 (2), 177-92.
Melkers, J. (2006), 'On the Road to Improved Performance', *Public Performance & Management Review,* 30 (1), 73 - 95.
Mises, Ludwig von (1949), *Human action : a treatise on economics* (London: London : Hodge).
Morgan, Gareth and Smircich, Linda (1980), 'The Case for Qualitative Research', *The Academy of Management Review,* 5 (4), 491-500.
Moynihan, Donald P. (2008), *The dynamics of performance management : constructing information and reform / Donald P. Moynihan* (Washington, D.C.: Georgetown University Press) xiii, 250.

Newcomer, K.E. (2007), 'How does program performance assessment affect program management in the federal government?', *Public Performance and Management Review,* 30 (3), 332-50.
Noordegraaf, Mirko and Abma, Tineke (2003), 'Management by Measurement? Public Management Practices Amidst Ambiguity', *Public Administration,* 81 (4), 853-71.
Norman, Richard (2002), 'Managing through Measurement or Meaning? Lessons from Experience with New Zealand's Public Sector Performance Management Systems', *International Review of Administrative Sciences,* 68 (4), 619-28.
North, Douglass C., et al. (2009), *Violence and social orders : a conceptual framework for interpreting recorded human history* (Cambridge: Cambridge University Press) xvii, 308 p.
Nove, Alec (1965), *The Soviet economy : an introduction* (2nd rev. edn., Minerva series of students' handbooks; London: Allen & Unwin) 356 p.
O'Toole, L. J. (1997), 'Treating networks seriously: Practical and research-based agendas in public administration', *Public Administration Review,* 57 (1), 45-52.
Olson, Mancur (1971), *The logic of collective action : public goods and the theory of groups / Mancur Olson* (New edn., Harvard economic studies ;; Cambridge, Mass ; London: Harvard University Press) x, 186.
Osborne, David and Gaebler, Ted (1993), *Reinventing government : how the entrepreneurial spirit is transforming the public sector / David Osborne and Ted Gaebler* (New York: Plume) xiii, 405.
Osborne, Stephen P. (2009), *The New Public Governance?: Emerging Perspectives on the Theory and Practice of Public Governance* (Taylor & Francis).
--- (2010), *The new public governance? emerging perspectives on the theory and practice of public governance / edited by Stephen P. Osborne* (London ; New York: Routledge) 1 online resource (xv, 431).
Ostrom, E. (1990), *Governing the commons: The evolution of institutions for collective action* (New York: Cambridge University Press).
--- (1998), 'A behavioral approach to the rational choice theory of collective action', *American Political Science Review,* 92 (1), 1-22.
--- (1999), 'Coping with tragedies of the commons', *Annual Review of Political Science,* 2, 493-535.
--- (2000), 'Collective action and the evolution of social norms', *Journal of Economic Perspectives,* 14 (3), 137-58.
--- (2004), 'Panarchy: understanding transformations in human and natural systems', *Ecological Economics,* 49, 488-91.
--- (2005), *A Grammar of Institutions* (Understanding Institutional Diversity) 137-+.
--- (2010), 'THE INSTITUTIONAL ANALYSIS AND DEVELOPMENT FRAMEWORK AND THE COMMONS', *Cornell Law Review,* 95 (4), 807-15.
Ostrom, E., et al. (1999), 'Sustainability - Revisiting the commons: Local lessons, global challenges', *Science,* 284 (5412), 278-82.
Otley, David (1999), 'Performance management: a framework for management control systems research', *Management Accounting Research,* 10 (4), 363-82.
Patrick, Rodger (2009), 'Performance management, gaming and police practice', (University of Birmingham).
Pidd, M. (2005), 'Perversity in public service performance measurement', *International Journal of Productivity and Performance Management,* 54 (5/6), 482-93.
Piekalkiewicz, Jaroslaw and Hamilton, Christopher (1991), *Public bureaucracies between reform and resistance : legacies, trends, and effects in china, the USSR, Poland, and Yugoslavia* (New York ; Oxford: Berg).
Piotrowski, S. J. and Rosenbloom, D. H. (2002), 'Nonmission-based values in results-oriented public management: The case of freedom of information', *Public Administration Review,* 62 (6), 643-57.
Piskotin, M. I. and Sayer, Jane (1989), *Socialism and state administration* (Moscow: Progress Publishers) 335p.
Pollitt, Christopher (2001), 'Convergence: The Useful Myth?', *Public Administration - Oxford,* 79 (4), 933-47.

--- (2006a), 'Performance management in practice: A comparative study of executive agencies', *Journal of Public Administration Research and Theory,* 16 (1), 25-44.
--- (2006b), 'Unleashing Change: A Study of Organizational Renewal in Government', *Administrative Science Quarterly,* 51 (2), 302-04.
Pollitt, Christopher and Bouckaert, Geert (2000), *Public management reform : a comparative analysis* (Oxford: Oxford UP) xiv, 314 p.
Pressman, Jeffrey L. and Wildavsky, Aaron B. (1984), *Implementation : how great expectations in Washington are dashed in Oakland : why it's amazing that federal programs work at all, this being a saga of the Economic Development Administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes / Jeffrey L. Pressman and Aaron Wildavsky* (3rd , expand edn.; London: University of California Press) xxvi, 281.
Propper, Carol and Wilson, Deborah (2003), 'The use and usefulness of performance measures in the public sector', *Oxford review of economic policy,* 19 (2), 250-67.
Public Administration Select Committee (2003), 'On Target? Government by Measurement', (London: The Stationery Office).
Radin, B (2008), 'The legacy of federal management change: PART repeats familiar

problems.', in F. Shea & T. Buss (ed.), *Performance management and budgeting: How

governments can learn from experience (Armonk, NY: M.E. Sharpe), 114–34.
Radnor, Zoe (2008), 'Muddled, massaging, manoeuvring or manipulated?', *International Journal of Productivity and Performance Management,* 57 (4), 316-28.
Radnor, Zoe and Barnes, David (2007), 'Historical analysis of performance measurement and management in operations management', *International Journal of Productivity and Performance Management,* 56 (5/6), 384-96.
Rangan, V. Kasturi (2004), 'Lofty Missions, Down-to-Earth Plans', *Harvard Business Review,* 82 (3), 112-19.
Ridgway, V. F. (1956), 'Dysfunctional Consequences of Performance Measurements', *Administrative Science Quarterly,* 1 (2), 240-47.
Rogers, Steve (1996), *Performance management in local government* (2nd edn.; London: Pitman) 144.
Rozin, P. and Royzman, E. B. (2001), 'Negativity bias, negativity dominance, and contagion', *Personality and Social Psychology Review,* 5 (4), 296-320.
Rubenstein, R., Schwartz, A. E., and Stiefel, L. (2003), 'Better than raw: A guide to measuring organizational performance with adjusted performance measures', *Public Administration Review,* 63 (5), 607-15.
Ryan, B. (2004), 'MEASURING AND MANAGING FOR PERFORMANCE: LESSONS FROM AUSTRALIA', in L. Jones, K. Schedler, and R. Mussari (eds.), *Strategies for Public Management Reform* (Research in Public Policy Analysis and Management, 13), 415-49.
Smith, Peter (1995), 'On the unintended consequences of publishing performance data in the public sector', *International Journal of Public Administration,* 18 (2), 277 - 310.
Talbot, Colin (2008), 'Performance Regimes—The Institutional Context of Performance Policies', *International Journal of Public Administration,* 31 (14), 1569-91.
ter Bogt, H. J. (2003), 'Performance evaluation styles in governmental organizations: How do professional managers facilitate politicians' work?', *Management Accounting Research,* 14 (4), 311-32.
Thiel, Sandra Van and Leeuw, Frans L. (2002), 'The Performance Paradox in the Public Sector', *Public Performance & Management Review,* 25 (3), 267-81.
Torres, L., Pina, V., and Yetano, A. (2011), 'Performance measurement in spanish local governments. A cross-case comparison study', *Public Administration,* 89 (3), 1081-109.
Tullock, Gordon (2005), *Bureaucracy* (Works. Selections. 2005; Indianapolis, Ind.: Lancaster : Liberty Fund ; Gazelle Drake Academic [distributor]) xvii, 437 p.

Vakkuri, Jarmo and Meklin, Pentti (2006), 'Ambiguity in Performance Measurement: A Theoretical Approach to Organisational Uses of Performance Measurement', *Financial Accountability & Management,* 22 (3), 235-50.

Verheijen, T. and Dobrolyubova, Y. (2007), 'Performance management in the Baltic States and Russia: success against the odds?', *International Review of Administrative Sciences,* 73, 205-15.

Weber, Max (1948), *From Max Weber : Essays in sociology / Translated, edited, and with an introd., by H.H. Gerth and C. Wright Mills*, ed. Hans Heinrich Gerth (London: Routledge & K. Paul) 490.

Willmott, Hugh (1993), 'Breaking the Paradigm Mentality', *Organization Studies,* 14 (5), 681-719.

Yakobson, L. (2006), '(Russian) Budjetnaya reforma: federalism ili upravlenie po resultatam (Budget reform: federalism vs performance management)', *Preprint WP8/2006/03, Gosudarstvennoe i municipalnoe upravlenie* (Moskva: Vyshaya Shkola Ekonomiki (Higher School of Economics)).

Yakovlev, A. (2010), 'Gosudarstvenny kapitalism, korrupziya i effectivnost' gosapparata (State capitalism, corruption and effectiveness of the public sector) ', *Obshestvennie nauki i sovremennost',* 4, 18-25.

Yang, K. F. (2011), 'THE SISYPHEAN FATE OF GOVERNMENT-WIDE PERFORMANCE ACCOUNTABILITY REFORMS Federal Performance Management Efforts and Employees' Daily Work, 2002-2008', *Public Performance and Management Review,* 35 (1), 149-76.

Yang, K. F. and Holzer, M. (2006), 'The performance-trust link: Implications for performance measurement', *Public Administration Review,* 66 (1), 114-26.

Zhigalov, D.V., Percov, L.V., and Chalaya, Y.Y. (2009), *Ispol'zovanie instrumentov budjetirovaniya, orientirovannogo na rezul'tat, v praktike municipalnogo upravleniya (The practice of using performance budgeting at the municipal level).* (Moscow: Fond "Institut ekonomiki goroda").

Zizek, S. (1989), *The Sublime object of ideology* (London. New York: Verso).

Appendix 8 Quotes in original Russian

[i] …в течение 2-х лет, это 7-8 годы, получили 20-ку базовых наших регионов, которая практически не менялась. То есть мы понимали, что это наши промышленные центры, это нефтегазовые регионы, которые имели серьёзную экономическую и, естественно, социальную базу.

[ii] Поэтому через 2 года работы этой системы мы перешли к другому показателю эффективности, это динамика показателей. Скажу, что в этот момент, это 2009 год, подверглись наибольшей критике, потому что мы полностью на 180 градусов перевернули всю систему и самый высокий рейтинг был у регионов, которые имели самую низкую базу, но в то же время по счётам показателей, динамики показателей, имели самые высокие показатели по динамике. То есть всё, что касается регионов Северного Кавказа, ряда наших не совсем базовых регионов Центральной России, они вышли на первое место в рейтинге.

[iii] Поэтому следующий шаг по совершенствованию самой методики – это уже 10-ый текущий год, это комплексная оценка, которая нивелировала уже и недостатки одной той системы базовой и системы, которая основывалась на динамике показателей. И уже с текущего года 2011, я думаю, что эта работа будет продолжена.

[iv] Интервьюер: А вот вы говорили, что губернаторов сравнивают на основе этих показателей

Респондент: Это игры, которые отвлекают, на самом деле.

[v] мы, как бы, спорим, с вами, хотя понимаем, что здесь для вас проблема, я думаю, в том, чтоб вам не навесили это в Министерство лесного хозяйства.

[vi] …есть отдельные показатели, за которые у нас отвечает каждый отдельный департамент, а в департаменте уже каждая конкретная проблема. У нашего управления есть показатели объема инвестиций и объема бюджетных инвестиций.

[vii] В докладе нашем (ДРОНД – прим. АСК) у нас есть отдельный раздел. Цель, по-моему, 4. Задача у нас, по-моему, 5. Основные задачи у моего управления, где я работаю, это доля инвалидов, получивших реабилитационные услуги от общего количества населения области. Вот это один показатель. У нас есть еще несколько показателей, которые уже относятся к задачам.

[viii] Чей это показатель? Т.е., с одной стороны, влияет на инвестиционную привлекательность, за которое отвечает наше министерство, а с другой стороны – эти показатели не наши…

[ix] …на самом деле, губернатор должен был дать задание Минстрою и сказать: ребята, вот вы, Минстрой, дайте мне четкую правильную пояснительную записку, где мы буксуем, какая у нас проблема. Давайте соберем рабочую группу и мы выясним, где эти проблемы. Но эта работа не была сделана.

[x] У нас есть, допустим, показатель полной частичной суммарной реабилитации инвалидов. Он зависит не только от нас. Он зависит от того, трудоустроен ли человек. Он зависит от здравоохранения, насколько его хорошо пролечили, что он мог выйти на работу. И от нас, какие реабилитационные услуги оказали… Возможно, под такими показателями должны стоять несколько министерств, которые разделят их между собой.

Интервьюер: А сейчас такого не происходит?

Респондент: Сейчас каждый отвечает за себя.

[xi] И, может быть, как раз в указе президента предусмотреть услуги, которые не только от одного ведомства зависят, а зависят от нескольких. Может, и губернатора включить в эти услуги. **Но кто возьмется?**

[xii] Ну, мы не можем ответить за трудоустройство. Хоть мы с ними общаемся и трудимся, и общие у нас мероприятия. Но это, межведомственное вот это взаимодействие, оно все-таки… (вздыхает).

[xiii] Все эти рейтинги – это ерунда, на самом деле.

[xiv] Потому что в результате они носят демонстрационный характер. И даже там, конечно, приятно получить какой-то грант, когда такие разные к исполнителям проектов. И они тебя похвалят, и, в общем-то, приятно в этом есть. Но, в общем, к этому относятся как к какой-то надстроечке, к процессу, потому что реально, ради получения грантов никто всерьез там не предпринимает больших усилий, чтобы изменились показатели там 825 или еще какие-нибудь. Ну, в общем, это как игра не стоит свеч. Нужно вложить очень много усилий, чтобы вот, их вкладывают, но не ради того, чтобы получить грант, они, ну мотивация другие. Гранты, в этом смысле, носят символический характер, а не мотивирующий

[xv] ...смотрит действительно, но о своих выводах необязательно кому-то рассказывать.

[xvi] "Отсутствие отлаженной процедуры распределения и взаимоувязки расходов с целями, задачами и программами приводит к тому, что субъекты бюджетного планирования по-прежнему сначала в течение года осуществляют деятельность в рамках своей компетенции, а затем уже в ходе подготовки ДРОНДов стараются вписать выполненные мероприятия в определенные докладом цели, задачи и показатели. Очевидно, что организованное таким образом бюджетирование как способ финансового планирования малоэффективно."

[xvii] Следует обратить внимание и на следующий нюанс: сложилась практика, когда подготовкой ДРОНДа занимается узкий круг сотрудников главного распорядителя бюджетных средств, а остальные сотрудники, включая руководителей других подразделений и руководителей подведомственных бюджетных учреждений, почти не принимают участия в данной работе. В такой ситуации текст ДРОНДа не имеет реальной связи с теми ориентирами, к достижению которых стремится персонал соответствующего ведомства в своей практической деятельности. Тем самым ориентация на конечный результат осуществляется лишь формально.

[xviii] М2: Они декларируемы. И как бы уровень их декларируемости не снижается. Они по-прежнему, но при этом как бы желание подробно контролировать все равно сохраняется, ну, во всяком случае, здесь на федерально уровне, конечном. Мне кажется, что вся проблема даже здесь не риторическая, вот, мне кажется, проблема скорее социально-экономическая стоит в стратегическом планировании. … Да и тогда, в общем, тоже вроде говорились такие же слова про диверсификацию экономики, там и так далее, а занимались-то все, в общем, совсем другой экономикой, которая там была связана с капитализацией нашего бизнеса за рубежом, нашего крупного бизнеса за рубеж… Это очень важно было, но это было как-то вроде мы так реально занимаемся этим, а декларируем цели другие. Конечно, важно тут жить в стране и дороги строить, больницы посещать, и так далее, так далее, так далее. При этом все так подразумевали, что то, чем мы занимаемся, мы не пишем.

[xix] А поскольку самая главная деятельность была за кадром, то вся остальная была более-менее декларативной.

[xx] Например, как можно повлиять на рождаемость? Вот здравоохранение может повлиять на рождаемость? У нас был один пример. Я рассказываю это в качестве анекдота, но это была правда жизни. Когда главный врач отчитывался в 1974 году, ему заместитель министра здравоохранения РСФСР сказал: «Что это у вас показатель снижается?» Он говорит: «Лично займусь». И на следующий год показатель вырос. Вы понимаете? Теперь ходит байка, что если главному врачу поручить, он любой показатель может исправить, в том числе и показатель рождаемости. [ГС-1]

[xxi] Зависимости от управленческих решений нет. За исключением одного решения — если нужно снизить заболеваемость, можно просто её год не регистрировать, и она упадёт. Конкретный год у нас будет вот такой (отличный - АСК). Смертность при определённых установках можно закопать: смертность туберкулезную в разные диагнозы записать и т.д. Вот туда списать пару сотен человек — тоже смертность упадет. Если так, то тогда этот показатель управляем. [ГС-1]

[xxii] Интервьюер: Обсуждаются ли эти показатели внутри организации?

Респондент: Нет. В последние годы нет. Когда только начинался ДРОНД, это обсуждалось. Я работал раньше в отрасли культуры. Там просто все встали на дыбы, потому что некоторые показатели проверить нельзя, т.е. нарисовать можно от балды. Кто это будет проверять? Сразу начали спрашивать: «А будут ли нас проверять?» Это первое. Потом так как-то замялось. Значит, не будут, значит, можно рисовать. Это было в первые годы. [ГС-3]

[xxiii] Часть показателей просто нельзя проверить. …Например, удельный вес населения, участвующих в культурно-досужный мероприятиях. Такой показатель был. Его никак не проверишь. Я даже сам не могу проверить. Я знаю, что районы, когда сдают отчётность, они это сами от балды рисуют. [ГС-3]

[xxiv] В 2008г запускали мониторинг ради этого ДРОНДа. И потом, когда все утряслось, когда люди поняли, что это можно рисовать, сидя у себя в деревне, все как по маслу пошло. В срок и с цифрами, какими надо. [ГС-3]

[xxv] Респондент: И поэтому библиотеки придерживали книжки, не списывали. Эти ленинские, марксистские тома специально они их копили ради того, чтобы выполнить показатели. Им надо новую литературу закупать, чтоб показатель выполнить. У них нет денег просто. Поэтому они всеми силами придерживают старье. Зато у них показатель идеальный.

Интервьюер: А почему они не рисовали показатель?

Респондент: Они боялись. Их можно проверить. А мероприятия то? Мероприятия можно по билетам проверить, удельный вес. А если это какие-то народные гуляния? Никак не проверишь. [ГС-3]

[xxvi] Я работал со статистикой, но в культуре. Статистику тоже рисуют, потому что статистика тоже проверяется. (по контексту имелось ввиду – «попадает в ДРОНД», а не «аудируется», АСК) У меня было количество клубных формирований, которое не должно падать. А клубные формирования в домах культуры. Дома культуры на селе остались только государственные. Я их учитывал. Они закрываются, потому что у них нет денег, да и вообще не целесообразно их там содержать. Соответственно, клубные формирования падают. Мне что с этим делать? Звонить и говорить: «Вы поняли, что нужно делать с клубным формированием?». Говорят: «Да, мы поняли».

Рисуют цифру. Потом стал вопрос — сколько мы так будем рисовать эти цифры? Тем более, динамика показателей должна расти. Мы до какого года будем её растить так искусственно?

Интервьюер: Т.е. они с каждым годом все дальше, дальше от жизни отходили, да?

Респондент: Да. И поэтому, когда с ДРОНДом всё замялось, все вздохнули с облегчением. [ГС-3]

[xxvii] Вот сейчас у нас новый губернатор, он хочет знать развёртку по области, что, как, в какой отрасли творится. Надо писать аналитические записки. Я сейчас звоню в культуру. Они работают там уже 2 с лишним года, и говорю: « Ну что там у вас с мониторингом, давайте посмотрим». «А у нас их никто не ведёт». Я говорю: «Как не ведёт?» **«Ну а как? ДРОНД не заставляют делать, мы и забросили».** [ГС-3]